

A Factor Graph Approach to Clock Offset Estimation in Wireless Sensor Networks

Aitzaz Ahmad, Davide Zennaro, Erchin Serpedin, and Lorenzo Vangelista

Abstract—The problem of clock offset estimation in a two way timing message exchange regime is considered when the likelihood function of the observation time stamps is Gaussian, exponential or log-normally distributed. A parametrized solution to the maximum likelihood (ML) estimation of clock offset, based on convex optimization, is presented, which differs from the earlier approaches where the likelihood function is maximized graphically. In order to capture the imperfections in node oscillators, which may render a time-varying nature to the clock offset, a novel Bayesian approach to the clock offset estimation is proposed by using a factor graph representation of the posterior density. Message passing using the max-product algorithm yields a closed form expression for the Bayesian inference problem. Several lower bounds on the variance of an estimator are derived for arbitrary exponential family distributed likelihood functions which, while serving as stepping stones to benchmark the performance of the proposed clock offset estimators, can be useful in their own right in classical as well Bayesian parameter estimation theory. To corroborate the theoretical findings, extensive simulation results are discussed for classical as well as Bayesian estimators in various scenarios. It is observed that the performance of the proposed estimators is fairly close to the fundamental limits established by the lower bounds.

Index Terms—Clock synchronization, factor graphs, message passing, estimation bounds, wireless sensor networks.

I. INTRODUCTION

WIRELESS sensor networks (WSNs) typically consist of a large number of geographically distributed sensor nodes, deployed to observe some phenomenon of interest. The nodes constituting such a network are low cost sensors that have limited abilities of data processing and communication. WSNs envisage tremendous applications in such diverse areas as industrial process control, battlefield surveillance, health monitoring, target localization and tracking, etc., [1]. With the recent advances in digital circuit technology, WSNs are expected to play a pivotal role in future wireless communications.

Clock synchronization in sensor networks is a critical component in data fusion and duty cycling operations, and has gained widespread interest in the past few years. Most of the current methods consider sensor networks exchanging time stamps based on the time at their respective clocks. A survey of the popular approaches employed in practice

for timing synchronization is presented in [2] and [3]. The *one-way* message exchange mechanism involves a reference node broadcasting its timing information to other nodes in a network. The receiver nodes record the arrival of these messages with respect to their own clock. After several such time stamps have been exchanged, the nodes estimate their offsets based on these observations. A particular case of this approach is the flooding time synchronization protocol (FTSP) [4] which uses regression to estimate the clock offset. On the other hand, through a *two-way* timing exchange process, adjacent nodes aim to achieve pairwise synchronization by communicating their timing information with each other. After a round of N messages, each node tries to estimate its own clock parameters. The timing-sync protocol for sensor networks (TPSNs) [5] uses this strategy in two phases to synchronize clocks in a network. The level discovery phase involves a spanning tree based representation of a WSN while nodes attempt to synchronize with their immediate parents using a two-way message exchange process in the synchronization phase. In *receiver-receiver* synchronization, nodes collect time stamps sent from a common broadcasting node and utilize them to adjust their clocks. The reference broadcast synchronization (RBS) protocol [6] uses reference beacons sent from a master node to establish a common notion of time across a network. An alternative framework for network-wide distributed clock synchronization consists of recasting the problem of agreement on oscillation phases and/or frequencies as a consensus based recursive model in which only local message passing is required among nodes. By assuming a connected network, it is possible to design efficient distributed algorithms by carefully choosing the update function. Under this framework, [9] proposed a Laplacian-based algorithm for establishing agreement on oscillation frequencies all over the network based on standard consensus. A combined agreement over both clock phases and frequencies has been studied in [10], by making use of state-of-the-art fast consensus techniques. Scalable synchronization algorithms for large sensor networks are developed in [11] and [12] inspired by mathematical biology models justifying synchrony in the biological agents.

The clock synchronization problem in a WSN offers a natural statistical signal processing framework whereby, the clock parameters are to be estimated using timing information from various sensors [22]. A model based synchronization approach to arrest the clock drifts is explored in [8]. The impairments in message transmission arise from the various delays experienced by the messages as they travel through the transmission medium. Therefore, a crucial component of

A. Ahmad and E. Serpedin are with the Department of Electrical and Computer Engineering, Texas A&M University, Texas, TX, 77843 USA.

D. Zennaro and L. Vangelista are with Department of Information Engineering, University of Padova, Padova, Italy.

A. Ahmad and E. Serpedin are supported by Qtel.

The work of D. Zennaro is partially supported by an “A. Gini” fellowship and has been performed while on leave at Texas A&M University, College Station, TX (USA).

efficient clock parameter estimation is accurate modeling of the network delay distributions. Several distributions have been proposed that aim to capture the random queuing delays in a network [7]. Some of these candidate distributions include exponential, Weibull, Gamma and log-normal distributions. Assuming an exponential delay distribution, several estimators were proposed in [13]. It was argued that when the propagation delay d is unknown, the maximum likelihood (ML) estimator for the clock offset θ is not unique. However, it was later shown in [14] that the ML estimator of θ does exist uniquely for the case of unknown d . The performance of these estimators was compared with benchmark estimation bounds in [15]. Considering an offset and skew model, Chaudhari *et.al.* presented algorithms for the joint ML estimation of clock offset and skew in [16] when the network delays are exponentially distributed. Clock offset and skew estimators were determined in [17] based on the assumption that the network delays arise from the contribution of several independent processes and as such, were modeled as Gaussian. The convergence of distributed consensus time synchronization algorithms is investigated in [18] and [19], assuming a Gaussian delay between sensor nodes. More recently, the minimum variance unbiased estimator (MVUE) for the clock offset under an exponential delay model was proposed in [20]. The timing synchronization problem for the offset-only case was also recast as an instance of convex optimization in [21] for Weibull distributed network delays. A recent contribution [23] has investigated the feasibility of determining the clock parameters by studying the fundamental limits on clock synchronization for wireline and wireless networks.

In this work, considering the classic two-way message exchange mechanism, a unified framework for the clock offset estimation problem is presented when the likelihood function of the observations is Gaussian, exponential or log-normally distributed. A parameterized solution is proposed for the ML estimation of clock offset by recasting the likelihood maximization as an instance of convex optimization. In order to incorporate the effect of time variations in the clock offset between sensor nodes, a Bayesian inference approach is also studied based on a factor graph representation of the posterior density. The major contributions of this work can be summarized as follows.

- 1) A unified framework for ML estimation of clock offset, based on convex optimization, is presented when the likelihood function of the observations is Gaussian, exponential and log-normally distributed. The proposed framework recovers the already known results for Gaussian and exponentially distributed likelihood functions and determines the ML estimate in case of log-normal distribution. Hence, the proposed convex optimization based approach represents a simpler alternative, and a more general derivation of ML estimator, which bypasses the graphical analysis used in [14] to maximize the likelihood function.
- 2) In order to capture the time variations in clock offsets due to imperfect oscillators, a Bayesian framework is presented by considering the clock offset as a random

Gauss-Markov process. Bayesian inference is performed using factor graphs and the max-product algorithm. The message passing strategy yields a closed form solution for Gaussian, exponential and log-normally distributed likelihood functions. This extends the current literature to cases where the clock offset may not be deterministic, but is in fact a random process.

- 3) In order to evaluate the performance of the proposed estimators, classical as well as Bayesian bounds are derived for arbitrary exponential family distributed likelihood functions, which is a wide class and contains almost all distributions of interest. While these results aid in comparing various estimators in this work, they can be useful in their own right in classical and Bayesian estimation theory.

This paper is organized as follows. The system model is outlined in Section II. The ML estimation of clock offset based on convex optimization is proposed in Section III. The factor graph based inference algorithm for the synchronization problem in a Bayesian paradigm is detailed in Section IV and a closed form solution is obtained. Section V presents several theoretical lower bounds on the variance of an estimator evaluated in the classical as well as Bayesian regime. Simulation studies are discussed in Section VI which corroborate the earlier results. Finally, the paper is concluded in Section VII along with some directions for future research.

II. SYSTEM MODEL

The process of pairwise synchronization between two nodes S and R is illustrated in Fig. 1. At the j th message exchange, node S sends the information about its current time through a message including time stamp T_j^1 . Upon receipt of this message, Node R records the reception time T_j^2 according to its own time scale. The two-way timing message exchange process is completed when node R replies with a synchronization packet containing time stamps T_j^2 and T_j^3 which is received at time T_j^4 by node S with respect to its own clock. After N such messages have been exchanged between nodes S and R , node S is equipped with time stamps $\{T_j^1, T_j^2, T_j^3, T_j^4\}_{j=1}^N$. The impairments in the signaling mechanism occur due to a *fixed* propagation delay, which accounts for the time required by the message to travel through the transmission medium, and a *variable* network delay, that arises due to queuing delay experienced by the messages during transmission and reception. By assuming that the respective clocks of nodes S and R are related by $C_R(t) = \theta + C_S(t)$, the two-way timing message exchange model at the j th instant can be represented as

$$\begin{aligned} T_j^2 &= T_j^1 + d + \theta + X_j \\ T_j^4 &= T_j^3 + d - \theta + Y_j \end{aligned} \quad (1)$$

where d represents the propagation delay, assumed symmetric in both directions, and θ is offset of the clock at node R relative to the clock at node S . X_j and Y_j are the independent and identically distributed variable network delays. By

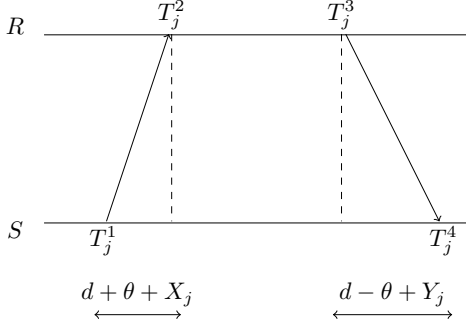


Fig. 1. A two-way timing message exchange mechanism

defining [13]

$$\begin{aligned} U_j &\triangleq T_j^2 - T_j^1 \\ V_j &\triangleq T_j^4 - T_j^3, \end{aligned}$$

the system in (1) can be equivalently expressed as

$$\begin{aligned} U_j &= d + \theta + X_j \\ V_j &= d - \theta + Y_j. \end{aligned} \quad (2)$$

By further defining

$$\xi \triangleq d + \theta \quad \psi \triangleq d - \theta, \quad (3)$$

the model in (2) can be written as

$$\begin{aligned} U_j &= \xi + X_j \\ V_j &= \psi + Y_j \end{aligned}$$

for $j = 1, \dots, N$. The goal is to determine precise estimates of ξ and ψ using observations $\{U_j, V_j\}_{j=1}^N$. An estimate of θ can, in turn, be obtained using (3) as follows

$$\theta = \frac{\xi - \psi}{2}. \quad (4)$$

Accurate modeling of the variable delays, X_j and Y_j , has been a topic of interest in recent years. Several distributions have been proposed that aim to capture the random effects caused by the queuing delays [7]. These distributions include exponential, gamma, log-normal and Weibull. In addition, the authors in [17] argued that X_j and Y_j result from contributions of numerous independent random processes and can, therefore, be assumed to be Gaussian. The ML estimate of d and θ for the case of exponential distribution was determined in [14]. Recently, the minimum variance unbiased estimate (MVUE) of the clock offset under an exponentially distributed network delay was proposed in [20]. In this work, instead of working with a specific distribution, a general framework of the clock synchronization problem is proposed that yields a parameterized solution of the clock offset estimation problem in the classical as well Bayesian regime when the likelihood function of the observations, U_j and V_j , is Gaussian, exponential or log-normally distributed.

In particular, the general notation used when the likelihood function of the observations $\mathbf{U} \triangleq [U_1, \dots, U_N]^T$ and

$\mathbf{V} \triangleq [V_1, \dots, V_N]^T$ is Gaussian or log-normally distributed is given below.

Unconstrained Likelihood:

$$f(\mathbf{U}; \xi) \propto \exp \left(\xi \sum_{j=1}^N \eta_\xi(U_j) - N\phi_\xi(\xi) \right) \quad (5)$$

$$f(\mathbf{V}; \psi) \propto \exp \left(\psi \sum_{j=1}^N \eta_\psi(V_j) - N\phi_\psi(\psi) \right) \quad (6)$$

where $\eta_\xi(U_j)$ and $\eta_\psi(V_j)$ are sufficient statistics for estimating ξ and ψ , respectively. The log-partition functions $\phi_\xi(\cdot)$ and $\phi_\psi(\cdot)$ serve as normalization factors so that $f(\mathbf{U}; \xi)$ and $f(\mathbf{V}; \psi)$ are valid probability distributions. The likelihood function is called ‘unconstrained’ since its domain is independent of the parameters ξ and ψ .

Similarly, the general notation used for an exponentially distributed likelihood function is given below.

Constrained Likelihood:

$$f(\mathbf{U}; \xi) \propto \exp \left(\xi \sum_{j=1}^N \eta_\xi(U_j) - N\phi_\xi(\xi) \right) \prod_{j=1}^N \mathbb{I}(U_j - \xi) \quad (7)$$

$$f(\mathbf{V}; \psi) \propto \exp \left(\psi \sum_{j=1}^N \eta_\psi(V_j) - N\phi_\psi(\psi) \right) \prod_{j=1}^N \mathbb{I}(V_j - \psi) \quad (8)$$

where the indicator function $\mathbb{I}(\cdot)$ is defined as

$$\mathbb{I}(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}.$$

and the roles of $\eta_\xi(U_j)$, $\eta_\psi(V_j)$, $\phi_\xi(\cdot)$ and $\phi_\psi(\cdot)$ are similar to (5) and (6). The likelihood function is called constrained since its domain depends on the parameters ξ and ψ . It must be noted that the likelihood functions (5)-(8) are expressed in terms of general exponential family distributions. This approach helps to keep the exposition sufficiently general and also allows us to recover the known results for the ML estimation of clock offset for Gaussian and exponentially distributed likelihood functions [14] [15], and determine the ML estimator of the clock offset in case of log-normally distributed likelihood function, as shown in Section III. The proposed approach will also prove useful in investigating a unified novel framework for clock offset estimation in the Bayesian setting for Gaussian, exponential or log-normally distributed likelihood functions, as will be shown in Section IV.

Some key ingredients of the proposed solution for the clock offset estimation problem, based on the properties of exponential family, can be summarized as follows [25].

- 1) The mean and variance of the sufficient statistic $\eta_\xi(U_j)$ are expressed as

$$\mathbb{E}[\eta_\xi(U_j)] = \frac{\partial \phi_\xi(\xi)}{\partial \xi} \quad (9)$$

$$\sigma_{\eta_\xi}^2 \triangleq \text{Var}[\eta_\xi(U_j)] = \frac{\partial^2 \phi_\xi(\xi)}{\partial \xi^2}. \quad (10)$$

- 2) The moment generating function (MGF) of the statistic $\eta_\xi(U_j)$ is given by

$$M_{\eta_\xi}(h) = \exp(\phi_\xi(\xi + h) - \phi_\xi(\xi)). \quad (11)$$

- 3) The non-negativity of the variance $\sigma_{\eta_\xi}^2$ in (10) implies that the log-partition function $\phi_\xi(\cdot)$ is convex.
 4) For Gaussian, exponential and log-normally distributed likelihood functions, the log-partition function $\phi_\xi(\xi)$ can be expressed as a second degree polynomial given by

$$\phi_\xi(\xi) = a_\xi \xi^2. \quad (12)$$

The coefficient a_ξ in this approximation can be obtained using the variance of the statistic $\eta_\xi(U_j)$, which is assumed known. Using (10), a_ξ is given by

$$a_\xi = \frac{\sigma_{\eta_\xi}^2}{2}.$$

If the statistical moment in (10) is not available, the empirical moment can be substituted since it readily follows from the weak law of large numbers that

$$\frac{\sum_{j=1}^N \eta_\xi(U_j)}{N} \xrightarrow{p} \mathbb{E}[\eta_\xi(U)], \quad N \rightarrow \infty$$

$$\frac{\sum_{j=1}^N \eta_\xi^2(U_j)}{N} \xrightarrow{p} \mathbb{E}[\eta_\xi^2(U)], \quad N \rightarrow \infty.$$

Similar expressions can also be written for $\eta_\psi(V_j)$, $M_{\eta_\psi}(h)$ and $\phi_\psi(\psi)$, respectively. Using the aforementioned properties of the exponential family, the ML as well as Bayesian estimates of ξ and ψ are to be determined utilizing the data set $\{U_j, V_j\}_{j=1}^N$, based on (12).

III. MAXIMUM LIKELIHOOD ESTIMATION

In this section, the ML estimates of θ are obtained by recasting the likelihood maximization as an instance of convex optimization. This approach differs from the graphical arguments used to maximize the likelihood in [14]. The specific cases of unconstrained and constrained likelihood functions are considered separately.

A. Unconstrained Likelihood

Using (5), (6) and (12), the unconstrained likelihood functions are given by

$$f(\mathbf{U}; \xi) \propto \exp\left(\xi \sum_{j=1}^N \eta_\xi(U_j) - N \frac{\sigma_{\eta_\xi}^2}{2} \xi^2\right) \quad (13)$$

$$f(\mathbf{V}; \psi) \propto \exp\left(\psi \sum_{j=1}^N \eta_\psi(V_j) - N \frac{\sigma_{\eta_\psi}^2}{2} \psi^2\right). \quad (14)$$

The ML estimates of ξ and ψ can now be expressed as

$$\hat{\xi}_{\text{ML}} = \arg \max_{\xi} \exp\left(\xi \sum_{j=1}^N \eta_\xi(U_j) - N \frac{\sigma_{\eta_\xi}^2}{2} \xi^2\right) \quad (15)$$

$$\hat{\psi}_{\text{ML}} = \arg \max_{\psi} \exp\left(\psi \sum_{j=1}^N \eta_\psi(V_j) - N \frac{\sigma_{\eta_\psi}^2}{2} \psi^2\right). \quad (16)$$

Theorem 1: The likelihood maximization problems (15) and (16) are strictly concave and the ML estimates are given by

$$\hat{\xi}_{\text{ML}} = \frac{\sum_{j=1}^N \eta_\xi(U_j)}{N \sigma_{\eta_\xi}^2} \quad (17a)$$

$$\hat{\psi}_{\text{ML}} = \frac{\sum_{j=1}^N \eta_\psi(V_j)}{N \sigma_{\eta_\psi}^2}. \quad (17b)$$

Hence, the ML estimator $\hat{\theta}_{\text{ML}}$ for the clock offset can be written as

$$\hat{\theta}_{\text{ML}} = \frac{\hat{\xi}_{\text{ML}} - \hat{\psi}_{\text{ML}}}{2}. \quad (18)$$

Proof: The ML estimate of ξ in (15) can be equivalently determined by maximizing the exponent in the likelihood function. It can be easily verified that the exponent in (15), which is a quadratic function of ξ , is strictly concave [29]. A similar explanation applies to the ML estimate of ψ . The ML estimates in (17) can be obtained by setting the first derivative of the exponent with respect to ξ (resp. ψ) to zero. The estimate $\hat{\theta}_{\text{ML}}$ in (18) can be obtained by invoking the invariance principle [28]. Hence, the proof readily follows. ■

1) *Gaussian Distributed Likelihood Function:* A particular application of Theorem 1 is the case when the likelihood functions $f(U_j; \xi)$ and $f(V_j; \psi)$ have a Gaussian distribution i.e., $f(U_j; \xi) \sim \mathcal{N}(\xi, \sigma_\xi^2)$ and $f(V_j; \psi) \sim \mathcal{N}(\psi, \sigma_\psi^2)$ [15]. Hence, it follows that

$$f(\mathbf{U}; \xi) = \frac{1}{(2\pi\sigma_\xi^2)^{\frac{N}{2}}} \exp\left(-\frac{\sum_{j=1}^N (U_j - \xi)^2}{2\sigma_\xi^2}\right) \quad (19)$$

which can be rearranged as

$$f(\mathbf{U}; \xi) \propto \exp\left(\xi \frac{\sum_{j=1}^N U_j}{\sigma_\xi^2} - \frac{N}{2\sigma_\xi^2} \xi^2\right).$$

By comparing with (13), we have

$$\eta_\xi(U_j) = \frac{U_j}{\sigma_\xi^2}, \quad \sigma_{\eta_\xi}^2 = \frac{1}{\sigma_\xi^2} \quad (20)$$

and the ML estimate using (17a) is given by

$$\hat{\xi}_{\text{ML}} = \frac{\sum_{j=1}^N U_j}{N}. \quad (21)$$

By a similar reasoning, the ML estimate for ψ (17b) is given by

$$\hat{\psi}_{\text{ML}} = \frac{\sum_{j=1}^N V_j}{N}.$$

Using (18), the ML estimate for the offset θ can be expressed as

$$\hat{\theta}_{\text{ML}} = \frac{\sum_{j=1}^N (U_j - V_j)}{2N}. \quad (22)$$

The above estimate coincides exactly with the one reported in [15].

2) *Log-Normally Distributed Likelihood Function*: Another application of Theorem 1 is when the samples U_j and V_j are log-normally distributed. In this case, we have

$$f(\mathbf{U}; \xi) = \frac{1}{\sqrt{2\pi\sigma_\xi^2}} \prod_{j=1}^N U_j^{-1} \exp\left(-\frac{\sum_{j=1}^N (\log U_j - \xi)}{2\sigma_\xi^2}\right) \\ \propto \exp\left(\xi \frac{\sum_{j=1}^N \log U_j}{\sigma_\xi^2} - \frac{N}{2\sigma_\xi^2} \xi^2\right). \quad (23)$$

A comparison with (13) yields

$$\eta_\xi(U_j) = \frac{\log U_j}{\sigma_\xi^2}, \quad \sigma_{\eta_\xi}^2 = \frac{1}{\sigma_\xi^2}.$$

Following a similar line of reasoning,

$$\eta_\psi(V_j) = \frac{\log V_j}{\sigma_\psi^2}, \quad \sigma_{\eta_\psi}^2 = \frac{1}{\sigma_\psi^2}.$$

The ML estimator for θ can be obtained from (18) using (17a) and (17b), and is given by

$$\hat{\theta}_{\text{ML}} = \frac{\sum_{j=1}^N (\log U_j - \log V_j)}{N}. \quad (24)$$

B. Constrained Likelihood

Using (7), (8) and (12), the constrained likelihood functions are given by

$$f(\mathbf{U}; \xi) \propto \exp\left(\xi \sum_{j=1}^N \eta_\xi(U_j) - N \frac{\sigma_{\eta_\xi}^2}{2} \xi^2\right) \prod_{j=1}^N \mathbb{I}(U_j - \xi) \quad (25)$$

$$f(\mathbf{V}; \psi) \propto \exp\left(\psi \sum_{j=1}^N \eta_\psi(V_j) - N \frac{\sigma_{\eta_\psi}^2}{2} \psi^2\right) \prod_{j=1}^N \mathbb{I}(V_j - \psi). \quad (26)$$

The resulting ML estimates of ξ and ψ can be obtained as

$$\hat{\xi}_{\text{ML}} = \arg \max_{\xi} \exp\left(\xi \sum_{j=1}^N \eta_\xi(U_j) - N \frac{\sigma_{\eta_\xi}^2}{2} \xi^2\right) \\ \text{such that } U_j \geq \xi \quad (27)$$

$$\hat{\psi}_{\text{ML}} = \arg \max_{\psi} \exp\left(\psi \sum_{j=1}^N \eta_\psi(V_j) - N \frac{\sigma_{\eta_\psi}^2}{2} \psi^2\right) \\ \text{such that } V_j \geq \psi. \quad (28)$$

Theorem 2: The likelihood maximization problems (27) and (28) are strictly concave and the ML estimates can be expressed as

$$\hat{\xi}_{\text{ML}} = \min\left(\frac{\sum_{j=1}^N \eta_\xi(U_j)}{N \sigma_{\eta_\xi}^2}, U_{(1)}\right) \quad (29a)$$

$$\hat{\psi}_{\text{ML}} = \min\left(\frac{\sum_{j=1}^N \eta_\psi(V_j)}{N \sigma_{\eta_\psi}^2}, V_{(1)}\right) \quad (29b)$$

where $U_{(1)}$ and $V_{(1)}$ denote the first order statistics of the samples U_j and V_j , respectively. The ML estimator $\hat{\theta}_{\text{ML}}$ for the clock offset is given by

$$\hat{\theta}_{\text{ML}} = \frac{\hat{\xi}_{\text{ML}} - \hat{\psi}_{\text{ML}}}{2}. \quad (30)$$

Proof: By using arguments similar to Theorem 1 and noting that the N constrains $U_j \geq \xi$ (resp. $V_j \geq \psi$) are linear functions of ξ (resp. ψ), the proof of concavity readily follows. Also, the likelihood maximization problems (27) and (28) can be equivalently expressed as

$$\hat{\xi}_{\text{ML}} = \arg \max_{\xi} \exp\left(\xi \sum_{j=1}^N \eta_\xi(U_j) - N \frac{\sigma_{\eta_\xi}^2}{2} \xi^2\right) \\ \text{such that } U_{(1)} \geq \xi \\ \hat{\psi}_{\text{ML}} = \arg \max_{\psi} \exp\left(\psi \sum_{j=1}^N \eta_\psi(V_j) - N \frac{\sigma_{\eta_\psi}^2}{2} \psi^2\right) \\ \text{such that } V_{(1)} \geq \psi.$$

The unconstrained maximizer $\bar{\xi}$ of the objective function in (27) is given by (17a). If $\bar{\xi} \leq U_{(1)}$, then $\hat{\xi}_{\text{ML}} = \bar{\xi}$. On the other hand, if $U_{(1)} \leq \bar{\xi}$, then the ML estimate is given by $\hat{\xi}_{\text{ML}} = U_{(1)}$ using concavity of the objective function. Combining the two cases, the ML estimate in (29a) is obtained. A similar explanation applies to the ML estimate $\hat{\psi}_{\text{ML}}$ in (29b) and $\hat{\theta}_{\text{ML}}$ follows from the invariance principle. ■

1) *Exponentially Distributed Likelihood Function*: For the case when the likelihood functions are exponentially distributed, the density function of the samples U_j can be written as [14]

$$f(\mathbf{U}; \xi) = \lambda_\xi^N \exp\left(-\lambda_\xi \sum_{j=1}^N (U_j - \xi)\right) \mathbb{I}(U_{(1)} - \xi) \quad (31)$$

where λ_ξ^{-1} is the mean of the delays X_j . The density function can be rearranged as

$$f(\mathbf{U}; \xi) \propto \exp(N \lambda_\xi \xi).$$

Comparing the above formulation with (25),

$$\eta_\xi(U_j) = \lambda_\xi, \quad \sigma_{\eta_\xi}^2 = 0. \quad (32)$$

Using (29a), the ML estimate is given by

$$\hat{\xi}_{\text{ML}} = U_{(1)}. \quad (33)$$

Employing a similar reasoning,

$$\hat{\psi}_{\text{ML}} = V_{(1)}.$$

Using (30), the ML estimate of θ is given by

$$\hat{\theta}_{\text{ML}} = \frac{U_{(1)} - V_{(1)}}{2}, \quad (34)$$

which coincides exactly with the one reported in [14], where it is derived using graphical arguments.

Remark 1: The ML estimation method outlined above differs from the previous work [14] in that it is based on convex

optimization, while [14] maximized the likelihood graphically. Hence, this approach presents an alternative view of the ML estimation of the clock offset. It also allows us to determine the ML estimator of θ when the likelihood function is log-normally distributed. In addition, Theorem 2 will also be useful in Section IV where estimation of θ in the Bayesian regime is discussed.

IV. A FACTOR GRAPH APPROACH

The imperfections introduced by environmental conditions in the quartz oscillator in sensor nodes results in a time-varying clock offset between nodes in a WSN. To cater for such a temporal variation, in this section, a Bayesian approach to the clock synchronization problem is adopted by representing the a-posteriori density as a factor graph. The inference is performed on the resulting factor graph by message passing using max-product algorithm. To ensure completeness, a brief description of factor graphs and the max-product algorithm is provided below.

A factor graph is a *bipartite* graph that represents a factorization of a global function as a product of local functions called *factors*, each factor being dependent on a subset of variables. Factor graphs are often used to produce a graphical model depicting the various inter-dependencies between a collection of interacting variables. Each factor is represented by a factor node and each variable has an edge or a half-edge. An edge connects a particular variable to a factor node if and only if it is an argument of the factor expressed by the factor node [24].

Inference can be performed by passing messages (sometimes called beliefs) along the edges of a factor graph. In particular, *max-product* algorithm is used to compute the messages exchanged between variables and factor nodes. These messages can be summarized as follows

variable to factor node :

$$m_{x \rightarrow f}(x) = \prod_{h \in n(x) \setminus f} m_{h \rightarrow x}(x) \quad (35)$$

factor node to variable :

$$m_{f \rightarrow x}(x) = \max_{\setminus \{x\}} \left(f(Z) \prod_{z \in n(f) \setminus \{x\}} m_{z \rightarrow f}(z) \right) \quad (36)$$

where $Z = n(f)$ is the set of arguments of the local function f . The marginal distributions associated with each variable can be obtained by the product of all incoming messages on the variable.

In order to sufficiently capture the temporal variations, the parameters ξ and ψ are assumed to evolve through a Gauss-Markov process given by

$$\begin{aligned} \xi_k &= \xi_{k-1} + w_k \\ \psi_k &= \psi_{k-1} + v_k \quad \text{for } k = 1, \dots, N \end{aligned}$$

where w_k and v_k are *i.i.d* such that $w_k, v_k \sim \mathcal{N}(0, \sigma^2)$. The posterior pdf can be expressed as

$$\begin{aligned} f(\xi, \psi | U, V) &\propto f(\xi, \psi) f(U, V | \xi, \psi) \\ &= f(\xi_0) \prod_{k=1}^N f(\xi_k | \xi_{k-1}) f(\psi_0) \prod_{k=1}^N f(\psi_k | \psi_{k-1}) \\ &\quad \cdot \prod_{k=1}^N f(U_k | \xi_k) f(V_k | \psi_k) \end{aligned} \quad (37)$$

where uniform priors $f(\xi_0)$ and $f(\psi_0)$ are assumed. Define $\delta_{k-1}^k \triangleq f(\xi_k | \xi_{k-1}) \sim \mathcal{N}(\xi_{k-1}, \sigma^2)$, $\nu_{k-1}^k \triangleq f(\psi_k | \psi_{k-1}) \sim \mathcal{N}(\psi_{k-1}, \sigma^2)$, $f_k \triangleq f(U_k | \xi_k)$, $h_k \triangleq f(V_k | \psi_k)$, where the likelihood functions are given by

$$\begin{aligned} f(U_k | \xi_k) &\propto \exp \left(\xi_k \eta_\xi(U_k) - \frac{\sigma_{\eta_\xi}^2}{2} \xi_k^2 \right) \\ f(V_k | \psi_k) &\propto \exp \left(\psi_k \eta_\psi(V_k) - \frac{\sigma_{\eta_\psi}^2}{2} \psi_k^2 \right), \end{aligned} \quad (38)$$

based on (12). The resulting factor graph representation of the posterior pdf is shown in Fig. 2.

Remark 2: A few important observations of this representation can be summarized below.

- Notice that the substitution in (3) renders a cycle-free nature to the factor graph. Therefore, inference by message passing on such a factor graph is indeed optimal [24].
- The two factor graphs shown in Fig. 2 have a similar structure and hence, message computations will only be shown for the estimate $\hat{\xi}_N$. Clearly, similar expressions will apply to $\hat{\psi}_N$.

In addition, only the case of constrained likelihood will be considered, since the case of an unconstrained likelihood is subsumed, as will be shown shortly. The clock offset estimator $\hat{\theta}_N$ can be obtained from $\hat{\xi}_N$ and $\hat{\psi}_N$ using (4).

By defining $\alpha_{\xi,k} \triangleq -\frac{\sigma_{\eta_\xi,k}^2}{2}$ and $\beta_{\xi,k} \triangleq \eta_\xi(U_k)$, the constrained likelihood function for the samples U_k can be written as

$$f_k \propto \exp(\alpha_{\xi,k} \xi_k^2 + \beta_{\xi,k} \xi_k) \mathbb{I}(U_k - \xi_k). \quad (39)$$

The message passing strategy starts by sending a message from the factor node f_N to the variable ξ_N . The variable ξ_N relays this message to the factor node δ_{N-1}^N . The factor node computes the product of this message with the factor δ_{N-1}^N and sends the resulting message to the variable ξ_{N-1} after ‘summarizing’ over the variable ξ_N . In the max-product algorithm, a ‘max’ function is used as a summary propagation operator (cf. (36)). These messages are computed as

$$\begin{aligned} m_{f_N \rightarrow \xi_N} &= f_N \\ m_{\xi_N \rightarrow \delta_{N-1}^N} &= f_N \\ m_{\delta_{N-1}^N \rightarrow \xi_{N-1}} &\propto \max_{\xi_N} \delta_{N-1}^N \cdot m_{\xi_N \rightarrow \delta_{N-1}^N} \\ &= \max_{\xi_N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left(\frac{-(\xi_N - \xi_{N-1})^2}{2\sigma^2} \right) \\ &\quad \cdot \exp(\alpha_{\xi,N} \xi_N^2 + \beta_{\xi,N} \xi_N) \mathbb{I}(U_N - \xi_N) \end{aligned}$$

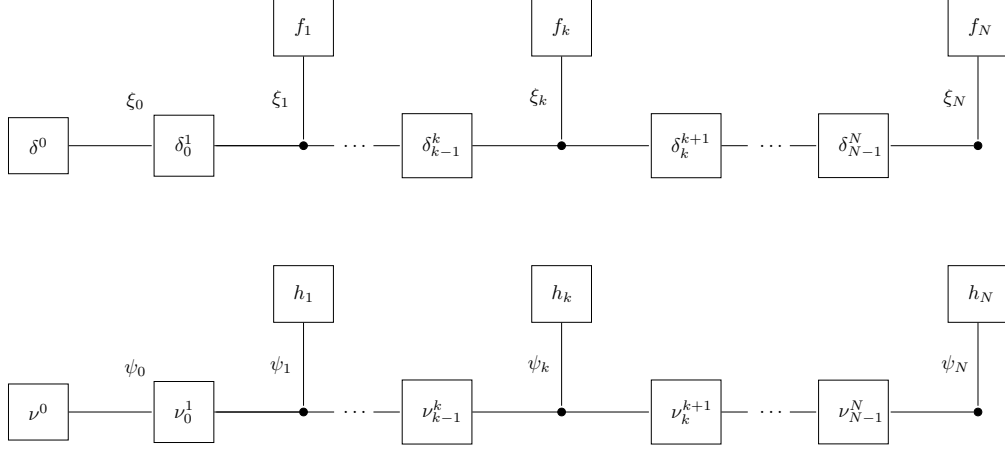


Fig. 2. Factor graph representation of the posterior density (37)

which can be rearranged as

$$m_{\delta_{N-1}^N \rightarrow \xi_{N-1}} \propto \max_{\xi_N \leq U_N} \exp(A_{\xi,N} \xi_N^2 + B_{\xi,N} \xi_{N-1}^2 + C_{\xi,N} \xi_N \xi_{N-1} + D_{\xi,N} \xi_N) \quad (40)$$

where

$$\begin{aligned} A_{\xi,N} &\triangleq -\frac{1}{2\sigma^2} + \alpha_{\xi,N}, & B_{\xi,N} &\triangleq -\frac{1}{2\sigma^2} \\ C_{\xi,N} &\triangleq \frac{1}{\sigma^2}, & D_{\xi,N} &\triangleq \beta_{\xi,N}. \end{aligned} \quad (41)$$

Let $\bar{\xi}_N$ be the unconstrained maximizer of the exponent in the objective function above, i.e.,

$$\bar{\xi}_N = \arg \max_{\xi_N} (A_{\xi,N} \xi_N^2 + B_{\xi,N} \xi_{N-1}^2 + C_{\xi,N} \xi_N \xi_{N-1} + D_{\xi,N} \xi_N).$$

This implies that

$$\bar{\xi}_N = -\frac{C_{\xi,N} \xi_{N-1} + D_{\xi,N}}{2A_{\xi,N}}. \quad (42)$$

Following a line of reasoning similar to Theorem 2, it follows that

$$\hat{\xi}_N = \min(\bar{\xi}_N, U_N).$$

However, $\bar{\xi}_N$ depends on ξ_{N-1} , which is undetermined at this stage. Hence, we need to further traverse the chain backwards. Assuming that $\bar{\xi}_N \leq U_N$, $\bar{\xi}_N$ from (42) can be plugged back in (40) which after some simplification yields

$$m_{\delta_{N-1}^N \rightarrow \xi_{N-1}} \propto \exp \left\{ \left(B_{\xi,N} - \frac{C_{\xi,N}^2}{4A_{\xi,N}} \right) \xi_{N-1}^2 - \frac{C_{\xi,N} D_{\xi,N}}{2A_{\xi,N}} \xi_{N-1} \right\}. \quad (43)$$

The message passed from the variable ξ_{N-1} to the factor node δ_{N-2}^{N-1} is the product of the message (43) and the message received from the factor node f_{N-1} , i.e.,

$$m_{\xi_{N-1} \rightarrow \delta_{N-2}^{N-1}} = m_{\delta_{N-1}^N \rightarrow \xi_{N-1}} \cdot m_{f_{N-1} \rightarrow \xi_{N-1}}.$$

Upon receipt of this message, the factor node δ_{N-2}^{N-1} delivers a product of this message and the factor δ_{N-2}^{N-1} to the variable node ξ_{N-2} after summarizing over ξ_{N-1} . This message can be expressed as

$$\begin{aligned} m_{\delta_{N-2}^{N-1} \rightarrow \xi_{N-2}} &\propto \max_{\xi_{N-1} \leq U_{N-1}} \delta_{N-2}^{N-1} \cdot m_{\xi_{N-1} \rightarrow \delta_{N-2}^{N-1}} \\ &= \max_{\xi_{N-1}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left(-\frac{(\xi_{N-1} - \xi_{N-2})^2}{2\sigma^2} \right) \\ &\quad \cdot \exp \left\{ \left(B_{\xi,N} - \frac{C_{\xi,N}^2}{4A_{\xi,N}} \right) \xi_{N-1}^2 - \frac{C_{\xi,N} D_{\xi,N}}{2A_{\xi,N}} \xi_{N-1} \right\} \\ &\quad \cdot \exp(\alpha_{\xi,N-1} \xi_{N-1}^2 + \beta_{\xi,N-1} \xi_{N-1}) \mathbb{I}(U_{N-1} - \xi_{N-1}). \end{aligned}$$

After some algebraic steps, the message above can be compactly represented as

$$m_{\delta_{N-2}^{N-1} \rightarrow \xi_{N-2}} \propto \max_{\xi_{N-1} \leq U_{N-1}} \exp(A_{\xi,N-1} \xi_{N-1}^2 + B_{\xi,N-1} \xi_{N-2}^2 + C_{\xi,N-1} \xi_{N-1} \xi_{N-2} + D_{\xi,N-1} \xi_{N-1}) \quad (44)$$

where

$$\begin{aligned} A_{\xi,N-1} &\triangleq -\frac{1}{2\sigma^2} + \alpha_{\xi,N-1} + B_{\xi,N} - \frac{C_{\xi,N}^2}{4A_{\xi,N}}, \\ B_{\xi,N-1} &\triangleq -\frac{1}{2\sigma^2}, & C_{\xi,N-1} &\triangleq \frac{1}{\sigma^2} \\ D_{\xi,N-1} &\triangleq \beta_{\xi,N-1} - \frac{C_{\xi,N} D_{\xi,N}}{2A_{\xi,N}}. \end{aligned}$$

Proceeding as before, the unconstrained maximizer $\bar{\xi}_{N-1}$ of the objective function above is given by

$$\bar{\xi}_{N-1} = -\frac{C_{\xi,N-1} \xi_{N-2} + D_{\xi,N-1}}{2A_{\xi,N-1}}$$

and the solution to the maximization problem (44) is expressed as

$$\hat{\xi}_{N-1} = \min(\bar{\xi}_{N-1}, U_{N-1}).$$

Again, $\bar{\xi}_{N-1}$ depends on ξ_{N-2} and therefore, the solution demands another traversal backwards on the factor graph

representation in Fig. 2. By plugging $\bar{\xi}_{N-1}$ back in (44), it follows that

$$m_{\delta_{N-2}^{N-1} \rightarrow \xi_{N-2}} \propto \exp \left\{ \left(B_{\xi, N-1} - \frac{C_{\xi, N-1}^2}{4A_{\xi, N-1}} \right) \xi_{N-2}^2 - \frac{C_{\xi, N-1} D_{\xi, N-1}}{2A_{\xi, N-1}} \xi_{N-2} \right\} \quad (45)$$

which has a form similar to (43). It is clear that one can keep traversing back in the graph yielding messages similar to (43) and (45). In general, for $i = 1, \dots, N-1$ we can write

$$\begin{aligned} A_{\xi, N-i} &\triangleq -\frac{1}{2\sigma^2} + \alpha_{\xi, N-i} + B_{\xi, N-i+1} - \frac{C_{\xi, N-i+1}^2}{4A_{\xi, N-i+1}} \\ B_{\xi, N-i} &\triangleq -\frac{1}{2\sigma^2}, \quad C_{\xi, N-i} \triangleq \frac{1}{\sigma^2} \\ D_{\xi, N-i} &\triangleq \beta_{\xi, N-i} - \frac{C_{\xi, N-i+1} D_{\xi, N-i+1}}{2A_{\xi, N-i+1}} \end{aligned} \quad (46)$$

and

$$\bar{\xi}_{N-i} = -\frac{C_{\xi, N-i} \xi_{N-i-1} + D_{\xi, N-i}}{2A_{\xi, N-i}} \quad (47)$$

$$\hat{\xi}_{N-i} = \min(\bar{\xi}_{N-i}, U_{N-i}). \quad (48)$$

Using (47) and (48) with $i = N-1$, it follows that

$$\bar{\xi}_1 = -\frac{C_{\xi, 1} \xi_0 + D_{\xi, 1}}{2A_{\xi, 1}} \quad (49)$$

$$\hat{\xi}_1 = \min(\bar{\xi}_1, U_1). \quad (50)$$

Similarly, by observing the form of (43) and (45), it follows that

$$m_{\delta_0^1 \rightarrow \xi_0} \propto \exp \left\{ \left(B_{\xi, 1} - \frac{C_{\xi, 1}^2}{4A_{\xi, 1}} \right) \xi_0^2 - \frac{C_{\xi, 1} D_{\xi, 1}}{2A_{\xi, 1}} \xi_0 \right\}. \quad (51)$$

The estimate $\hat{\xi}_0$ can be obtained by maximizing the received message in (51). It can be noticed from the structure of the factor graph that this maximization is inherently unconstrained i.e.,

$$\begin{aligned} \hat{\xi}_0 &= \bar{\xi}_0 = \max_{\xi_0} m_{\delta_0^1 \rightarrow \xi_0} \\ \Rightarrow \hat{\xi}_0 &= \frac{C_{\xi, 1} D_{\xi, 1}}{4A_{\xi, 1} B_{\xi, 1} - C_{\xi, 1}^2}. \end{aligned} \quad (52)$$

The estimate in (52) can now be substituted in (49) to yield $\bar{\xi}_1$, which can then be used to solve for $\hat{\xi}_1$ in (50). Clearly, this chain of calculations can be continued using recursions (47) and (48).

Define

$$g_{\xi, k}(x) \triangleq -\frac{C_{\xi, k} x + D_{\xi, k}}{2A_{\xi, k}}. \quad (53)$$

A key property of the function $g_{\xi, k}(\cdot)$, which proves useful in the quest for a closed form solution, can be summarized in the following lemma.

Lemma 1: For real numbers a and b , the function $g_{\xi, k}(\cdot)$ defined in (53) satisfies

$$g_{\xi, k}(\min(a, b)) = \min(g_{\xi, k}(a), g_{\xi, k}(b)).$$

Proof: The constants $A_{\xi, k}$, $C_{\xi, k}$ and $D_{\xi, k}$ are defined in (41) and (46). The proof follows by noting that $\frac{-C_{\xi, k}}{2A_{\xi, k}} > 0$ which implies that $g_{\xi, k}(\cdot)$ is a monotonically increasing function. ■

With the notation $g_{\xi, k}(\cdot)$, the following chain of equalities can be conveniently written as

$$\begin{aligned} \bar{\xi}_1 &= g_{\xi, 1}(\hat{\xi}_0) \\ \hat{\xi}_1 &= \min(U_1, g_{\xi, 1}(\hat{\xi}_0)) \\ \bar{\xi}_2 &= g_{\xi, 2}(\hat{\xi}_1) \\ \hat{\xi}_2 &= \min(U_2, g_{\xi, 2}(\hat{\xi}_1)) \end{aligned}$$

where

$$\begin{aligned} g_{\xi, 2}(\hat{\xi}_1) &= g_{\xi, 2}(\min(U_1, g_{\xi, 1}(\hat{\xi}_0))) \\ &= \min(g_{\xi, 2}(U_1), g_{\xi, 2}(g_{\xi, 1}(\hat{\xi}_0))) \end{aligned} \quad (54)$$

where (54) follows from Lemma 1. The estimate $\hat{\xi}_2$ can be expressed as

$$\begin{aligned} \hat{\xi}_2 &= \min(U_2, \min(g_{\xi, 2}(U_1), g_{\xi, 2}(g_{\xi, 1}(\hat{\xi}_0)))) \\ &= \min(U_2, g_{\xi, 2}(U_1), g_{\xi, 2}(g_{\xi, 1}(\hat{\xi}_0))). \end{aligned}$$

By following the same procedure, one can write

$$\begin{aligned} \hat{\xi}_3 &= \min(U_3, g_{\xi, 3}(U_2), g_{\xi, 3}(g_{\xi, 2}(U_1)), \\ &\quad g_{\xi, 3}(g_{\xi, 2}(g_{\xi, 1}(\hat{\xi}_0)))) \end{aligned}$$

For $m \geq j$, define

$$G_{\xi, j}^m(\cdot) \triangleq g_{\xi, m}(g_{\xi, m-1} \dots g_{\xi, j}(\cdot)). \quad (55)$$

The estimate $\hat{\xi}_3$ can, therefore, be compactly represented as

$$\hat{\xi}_3 = \min(U_3, G_{\xi, 3}^3(U_2), G_{\xi, 2}^3(U_1), G_{\xi, 1}^3(\hat{\xi}_0)).$$

Hence, one can keep estimating $\hat{\xi}_k$ at each stage using this strategy. Note that the estimator only depends on functions of data and can be readily evaluated.

In order to derive analogous expressions for ψ , a similar line of reasoning should be followed. In particular, the constants $A_{\xi, N-i}$, $B_{\xi, N-i}$, $C_{\xi, N-i}$ and $D_{\xi, N-i}$ for $i = 0, \dots, N-1$, can be obtained straightforwardly from (41) and (46) by substituting $\alpha_{\xi, N-i}$ and $\beta_{\xi, N-i}$ with $\alpha_{\psi, N-i}$ and $\beta_{\psi, N-i}$, respectively. Using these constants, $\hat{\psi}_0$, $g_{\psi, k}$ and $G_{\psi, j}^m$ can be defined analogously to (52), (53) and (55).

Generalizing this framework, the closed form expression for the clock offset estimate $\hat{\theta}_N$ is given by the following theorem.

Theorem 3: The state estimates $\hat{\xi}_N$ and $\hat{\psi}_N$ for the posterior pdf in (37) can be expressed as

$$\begin{aligned} \hat{\xi}_N &= \min(U_N, G_{\xi, N}^N(U_{N-1}), \dots, G_{\xi, 2}^N(U_1), G_{\xi, 1}^N(\hat{\xi}_0)) \\ \hat{\psi}_N &= \min(V_N, G_{\psi, N}^N(V_{N-1}), \dots, G_{\psi, 2}^N(V_1), G_{\psi, 1}^N(\hat{\psi}_0)) \end{aligned}$$

and the factor graph based clock offset estimate (FGE) $\hat{\theta}_N$ is given by

$$\hat{\theta}_N = \frac{\hat{\xi}_N - \hat{\psi}_N}{2}. \quad (56)$$

Proof: The proof follows from the discussion above and using (4). ■

Remark 3: The closed form expressions in Theorem 3 enable the estimation of the clock offset, when it may be time-varying and the likelihood functions, $f(U_k|\xi_k)$ and $f(V_k|\psi_k)$, have a Gaussian, exponential or log-normal distribution.

A. Gaussian Distributed Likelihood Function

A particular case of the Bayesian framework described above occurs when likelihood functions $f(U_k|\xi_k)$ and $f(V_k|\psi_k)$ have a Gaussian distribution, i.e., $f(U_k|\xi_k) \sim \mathcal{N}(\xi_k, \sigma_{\xi,k}^2)$ and $f(V_k|\psi_k) \sim \mathcal{N}(\psi_k, \sigma_{\psi,k}^2)$, i.e.,

$$\begin{aligned} f(U_k|\xi_k) &= \frac{1}{\sqrt{2\pi\sigma_{\xi,k}^2}} \exp\left\{-\frac{(U_k - \xi_k)^2}{2\sigma_{\xi,k}^2}\right\} \\ &\propto \exp\left(\frac{\xi_k U_k}{2\sigma_{\xi,k}^2} - \frac{\xi_k^2}{2\sigma_{\xi,k}^2}\right). \end{aligned} \quad (57)$$

The aforementioned Gaussian distribution constitutes an unconstrained likelihood function, i.e., the domain of the pdf is independent of the unknown parameter ξ_k . It is clear from the message passing approach that at each stage k of the factor graph, the unconstrained maximizer $\bar{\xi}_k$ is the actual solution to the likelihood maximization problem

$$\max_{\xi_k} \exp(A_{\xi,k}\xi_k^2 + B_{\xi,k}\xi_{k-1}^2 + C_{\xi,k}\xi_k\xi_{k-1} + D_{\xi,k}\xi_k)$$

i.e., $\hat{\xi}_k = \bar{\xi}_k \forall k = 1, \dots, N$. Hence, the unconstrained likelihood maximization problem is subsumed in the message passing framework for constrained likelihood maximization. It follows from Theorem 3 that $\hat{\xi}_N$ for Gaussian distributed observations U_k in (57) is given by

$$\hat{\xi}_N = G_{\xi,1}^N(\hat{\xi}_0)$$

where $\hat{\xi}_0$ and $G_{\xi,1}^N(\cdot)$ are defined in (52) and (55), respectively. Evaluating $\hat{\xi}_N$ requires to determine the constants in (41) and (46). By comparing (57) with (39), we have

$$\alpha_{\xi,k} = -\frac{1}{2\sigma_{\xi,k}^2}, \quad \beta_{\xi,k} = \frac{U_k}{\sigma_{\xi,k}^2}. \quad (58)$$

Using these values for $\alpha_{\xi,k}$ and $\beta_{\xi,k}$, (41) and (46) can be written as

$$\begin{aligned} A_{\xi,N} &= -\frac{1}{2\sigma^2} - \frac{1}{2\sigma_{\xi,N}^2}, \quad B_{\xi,N} = -\frac{1}{2\sigma^2} \\ C_{\xi,N} &= \frac{1}{\sigma^2}, \quad D_{\xi,N} = \frac{U_N}{\sigma_{\xi,N}^2} \\ A_{\xi,N-i} &= -\frac{1}{2\sigma^2} - \frac{1}{2\sigma_{\xi,N-i}^2} + B_{\xi,N-i+1} - \frac{C_{\xi,N-i+1}^2}{4A_{\xi,N-i+1}} \\ B_{\xi,N-i} &= -\frac{1}{2\sigma^2}, \quad C_{\xi,N-i} = \frac{1}{\sigma^2} \\ D_{\xi,N-i} &= \frac{U_{N-i}}{\sigma_{\xi,N-i}^2} - \frac{C_{\xi,N-i+1}D_{\xi,N-i+1}}{2A_{\xi,N-i+1}} \end{aligned} \quad (59)$$

for $i = 1, \dots, N-1$. Using similar arguments, it can be shown that the estimate $\hat{\psi}_N$ is given by

$$\hat{\psi}_N = G_{\psi,1}^N(\hat{\psi}_0).$$

It follows from (4) that the FGE, $\hat{\theta}_N$, can be expressed as

$$\hat{\theta}_N = \frac{G_{\xi,1}^N(\hat{\xi}_0) - G_{\psi,1}^N(\hat{\psi}_0)}{2}. \quad (60)$$

The behavior of $\hat{\theta}_N$ can be further investigated for the case when the noise variance σ^2 in the Gauss-Markov model goes to zero. Consider

$$g_{\xi,N}(\xi) = -\frac{C_{\xi,N}\xi + D_{\xi,N}}{2A_{\xi,N}}$$

where the constants $A_{\xi,N}$, $B_{\xi,N}$, $C_{\xi,N}$ and $D_{\xi,N}$ are given by (59). After some algebraic steps, we have

$$g_{\xi,N}(\xi) = \frac{\sigma_{\xi}^2\xi + \sigma^2U_N}{\sigma_{\xi}^2 + \sigma^2}.$$

As $\sigma^2 \rightarrow 0$, $g_{\xi,N}(\xi) \rightarrow \xi$. Similarly, it can be shown that $g_{\psi,N-1}(\xi) \rightarrow \xi$ as $\sigma^2 \rightarrow 0$. Hence, it follows that in the low system noise regime, as $\sigma^2 \rightarrow 0$

$$\hat{\xi}_N \rightarrow \hat{\xi}_0 = \frac{C_{\xi,1}D_{\xi,1}}{4A_{\xi,1}B_{\xi,1} - C_{\xi,1}^2}.$$

Similarly, it can be shown that

$$\hat{\psi}_N \rightarrow \hat{\psi}_0 = \frac{C_{\psi,1}D_{\psi,1}}{4A_{\psi,1}B_{\psi,1} - C_{\psi,1}^2}.$$

Therefore

$$\hat{\theta}_N \rightarrow \frac{\hat{\xi}_0 - \hat{\psi}_0}{2},$$

which can be proven to be equal to the ML estimator (22).

B. Log-Normally Distributed Likelihood Function

The log-normally distributed likelihood function in the Bayesian regime can be expressed as

$$\begin{aligned} f(U_k|\xi_k) &= \frac{1}{U_k\sigma_{\xi,k}\sqrt{2\pi}} \exp\left(-\frac{(\log U_k - \xi_k)^2}{2\sigma_{\xi,k}^2}\right) \\ &\propto \exp\left(\frac{\xi_k \log(U_k)}{2\sigma_{\xi,k}^2} - \frac{\xi_k^2}{2\sigma_{\xi,k}^2}\right). \end{aligned} \quad (61)$$

By comparing (61) and (39), we have

$$\alpha_{\xi,k} = -\frac{1}{2\sigma_{\xi,k}^2} \quad \beta_{\xi,k} = \frac{\log U_k}{\sigma_{\xi,k}^2}.$$

Clearly, the only difference here with the Gaussian distribution is a redefinition of $\beta_{\xi,k}$. The expression of $\hat{\xi}_N$ in this case is again

$$\hat{\xi}_N = G_{\xi,1}^N(\hat{\xi}_0)$$

where $G_{\xi,1}^N(\cdot)$ and $\hat{\xi}_0$ are given by (55) and (52), respectively. The recursively evaluated constants in (41) and (46) can be written as

$$\begin{aligned} A_{\xi,N} &= -\frac{1}{2\sigma^2} - \frac{1}{2\sigma_{\xi,N}^2}, \quad B_{\xi,N} = -\frac{1}{2\sigma^2} \\ C_{\xi,N} &= \frac{1}{\sigma^2}, \quad D_{\xi,N} = \frac{\log U_N}{\sigma_{\xi,N}^2} \\ A_{\xi,N-i} &= -\frac{1}{2\sigma^2} - \frac{1}{2\sigma_{\xi,N-i}^2} + B_{\xi,N-i+1} - \frac{C_{\xi,N-i+1}^2}{4A_{\xi,N-i+1}} \\ B_{\xi,N-i} &= -\frac{1}{2\sigma^2}, \quad C_{\xi,N-i} = \frac{1}{\sigma^2} \\ D_{\xi,N-i} &= \frac{\log U_{N-i}}{\sigma_{\xi,N-i}^2} - \frac{C_{\xi,N-i+1}D_{\xi,N-i+1}}{2A_{\xi,N-i+1}} \end{aligned}$$

for $i = 1, \dots, N-1$. Similar arguments apply to ψ . Hence, the FGE $\hat{\theta}_N$ can be expressed as

$$\hat{\theta}_N = \frac{G_{\xi,1}^N(\hat{\xi}_0) - G_{\psi,1}^N(\hat{\psi}_0)}{2}. \quad (62)$$

Again, as the Gauss-Markov system noise $\sigma^2 \rightarrow 0$, the above estimator approaches its ML counterpart (24).

C. Exponential Distribution

Theorem 3 can also be used to derive a Bayesian estimator $\hat{\xi}_N$ for the exponentially distributed likelihood case considered in [14]. In this case, we have

$$\begin{aligned} f(U_k|\xi_k) &= \lambda_\xi \exp(-\lambda_\xi(U_k - \xi_k)) \mathbb{I}(U_k - \xi_k) \\ &\propto \exp(\lambda_\xi \xi_k) \mathbb{I}(U_k - \xi_k) \end{aligned} \quad (63)$$

where λ_ξ^{-1} is the mean network delay of X_k . A comparison of (63) with (39) reveals that

$$\alpha_{\xi,k} = 0, \quad \beta_{\xi,k} = \lambda_\xi.$$

For these values of $\alpha_{\xi,k}$ and $\beta_{\xi,k}$, the constants $A_{\xi,k}$, $B_{\xi,k}$, $C_{\xi,k}$ and $D_{\xi,k}$ are given by

$$\begin{aligned} A_{\xi,k} &= -\frac{1}{2\sigma^2}, \quad B_{\xi,k} = -\frac{1}{2\sigma^2} \\ C_{\xi,k} &= \frac{1}{\sigma^2}, \quad D_{\xi,k} = \lambda_\xi \end{aligned}$$

for all $k = 1, \dots, N$. Using Theorem 3, we have

$$\begin{aligned} G_{\xi,N}^N(U_{N-1}) &= -\frac{C_{\xi,N}U_{N-1} + D_{\xi,N}}{2A_{\xi,N}} \\ &= U_{N-1} + \lambda_\xi \sigma^2. \end{aligned}$$

Similarly it can be shown that

$$G_{\xi,N-1}^N(U_{N-2}) = U_{N-2} + 2\lambda_\xi \sigma^2$$

and so on. The estimator $\hat{\xi}_0$ at the last step can be evaluated as

$$\hat{\xi}_0 = \frac{C_{\xi,1}D_{\xi,1}}{4A_{\xi,1}B_{\xi,1} - C_{\xi,1}^2} = +\infty.$$

This implies that

$$G_{\xi,1}^N(\hat{\xi}_0) = +\infty. \quad (64)$$

Using (64) and Theorem 3, it readily follows that

$$\begin{aligned} \hat{\xi}_N &= \min(U_N, U_{N-1} + \lambda_\xi \sigma^2, U_{N-2} + 2\lambda_\xi \sigma^2, \\ &\quad \dots, U_1 + (N-1)\lambda_\xi \sigma^2). \end{aligned} \quad (65)$$

Similarly, for the pdf [14]

$$\begin{aligned} f(V_k|\psi_k) &= \lambda_\psi \exp(-\lambda_\psi(V_k - \psi_k)) \mathbb{I}(V_k - \psi_k) \\ &\propto \exp(\lambda_\psi \psi_k) \mathbb{I}(V_k - \psi_k), \end{aligned}$$

the estimate $\hat{\psi}_N$ is given by

$$\begin{aligned} \hat{\psi}_N &= \min(V_N, V_{N-1} + \lambda_\psi \sigma^2, V_{N-2} + 2\lambda_\psi \sigma^2, \\ &\quad \dots, V_1 + (N-1)\lambda_\psi \sigma^2) \end{aligned} \quad (66)$$

and the estimate $\hat{\theta}_N$ can be obtained using (56), (65) and (66) as

$$\begin{aligned} \hat{\theta}_N &= \frac{1}{2} \min(U_N, U_{N-1} + \lambda_\xi \sigma^2, U_{N-2} + 2\lambda_\xi \sigma^2, \\ &\quad \dots, U_1 + (N-1)\lambda_\xi \sigma^2) - \\ &\quad \frac{1}{2} \min(V_N, V_{N-1} + \lambda_\psi \sigma^2, V_{N-2} + 2\lambda_\psi \sigma^2, \\ &\quad \dots, V_1 + (N-1)\lambda_\psi \sigma^2). \end{aligned} \quad (67)$$

As the Gauss-Markov system noise $\sigma^2 \rightarrow 0$, (67) yields

$$\hat{\theta}_N \rightarrow \hat{\theta}_{\text{ML}} = \frac{\min(U_N, \dots, U_1) - \min(V_N, \dots, V_1)}{2},$$

which is the ML estimator given by (34).

V. CLASSICAL AND BAYESIAN BOUNDS

To evaluate the performance of the estimators derived in the preceding sections, classical as well as Bayesian lower bounds on the variance of the estimators are discussed. The placement of a lower bound allows one to compare estimators by plotting their performance against the bound. It must be emphasized here that the results in this section assume no specific form of the log-partition function and are therefore, valid for arbitrary distributions from the exponential family, which is a wide class and contains almost all distributions of interest. Hence, these results are fairly general and can be useful in their own right in classical as well as Bayesian parameter estimation theory, and at the same time will be used as a stepping stone towards comparing the estimators developed thus far.

The likelihood function of the data is considered an arbitrary member of the exponential family of distributions. In addition, depending on whether the domain of the likelihood depends on the parameter to be estimated, both cases of unconstrained as well as constrained likelihood functions are discussed to maintain full generality. The general expressions for the unconstrained and constrained likelihood functions for observations $\mathbf{Z} \triangleq [Z_1, \dots, Z_N]^T$ are given by

Unconstrained Likelihood:

$$f(\mathbf{Z}; \rho) \propto \exp \left(\rho \sum_{j=1}^N \eta(Z_j) - N\phi(\rho) \right) \quad (68)$$

Constrained Likelihood:

$$f(\mathbf{Z}; \rho) \propto \exp \left(\rho \sum_{j=1}^N \eta(Z_j) - N\phi(\rho) \right) \prod_{j=1}^N \mathbb{I}(Z_j - \rho) \quad (69)$$

where ρ is the scalar parameter to be estimated. The goal is to derive lower bounds on the variance of estimators of ρ . For the case of classical estimation, the Cramer-Rao and Chapman-Robbins bounds are considered, while the Bayesian Cramer-Rao bound and a Bayesian version of the Chapman-Robbins bound are derived for the Bayesian paradigm.

A. Cramer-Rao Lower Bound

The Cramer-Rao lower bound (CRB) is a lower bound on the variance of an unbiased estimator of a deterministic parameter. It is useful primarily because it is relatively simple to compute. However, it relies on certain ‘regularity conditions’ which are not satisfied by constrained likelihood functions when the domain of the likelihood depends on the unknown parameter (cf. (69)). Hence, CRB is determined for the case of unconstrained likelihood functions only.

In particular, CRB states that the variance of an unbiased estimator of ρ is lower bounded by

$$\text{Var}(\hat{\rho}) \geq \frac{-1}{\mathbb{E} \left[\frac{\partial^2 \ln f(\mathbf{Z}; \rho)}{\partial \rho^2} \right]} \quad (70)$$

Theorem 4: The CRB for ρ in the unconstrained likelihood function in (68) is given by

$$\text{Var}(\hat{\rho}) \geq \frac{1}{N\sigma_\eta^2} \quad (71)$$

where

$$\sigma_\eta^2 = \frac{\partial^2 \phi(\rho)}{\partial \rho^2}.$$

Proof: The Fisher information for the likelihood function is given by

$$\begin{aligned} I(\rho) &\triangleq \mathbb{E} \left[\frac{\partial^2 \ln f(\mathbf{Z}; \rho)}{\partial \rho^2} \right] \\ &= -N \frac{\partial^2 \phi(\rho)}{\partial \rho^2} = -N\sigma_\eta^2 \end{aligned}$$

and the proof readily follows. ■

B. Chapman-Robbins Bound

Chapman-Robbins bound (CHRB), proposed in [26], sets a lower bound on the variance of an estimator of a deterministic parameter. The CHRB does not make any assumptions on the differentiability of the likelihood function and the regularity conditions that often constrain the use of CRB, and is substantially tighter than the CRB in many situations. Hence, CHRB

is employed to determine a lower bound on the variance of an unbiased estimator of ρ for constrained likelihood functions.

In general for a parameter ρ , the CHRB is given by

$$\text{Var}(\hat{\rho}) \geq \left[\inf_h \frac{1}{h^2} \left\{ \mathbb{E} \left(\frac{f(\mathbf{Z}; \rho + h)}{f(\mathbf{Z}; \rho)} \right)^2 - 1 \right\} \right]^{-1}, \quad (72)$$

which can be evaluated as shown below.

Theorem 5: The CHRB for the parameter ρ given the likelihood function (69) can be expressed as

$$\text{Var}(\hat{\rho}) \geq \left[\inf_h \frac{\left\{ (M_\eta(h))^{-2N} \cdot \zeta^N(h) - 1 \right\}}{h^2} \right]^{-1} \quad (73)$$

where $M_\eta(h)$ is the MGF of the statistic $\eta(Z_j)$ and

$$\zeta(h) \triangleq \mathbb{E} [\exp(2h\eta(Z_j)) \mathbb{I}(Z_j - \rho - h)] \quad (74)$$

with the expectation taken with respect to any Z_j .

Proof: The details of the proof are relegated to Appendix A. ■

C. Bayesian Cramer-Rao Lower Bound

The Bayesian Cramer-Rao bound (BCRB) is a lower bound on the variance of an unbiased estimator when the parameter assumes a prior density. It requires the same regularity conditions to be satisfied as its classical counterpart.

For an estimator $\hat{\rho}_k$ of ρ_k , the BCRB states that the variance of the estimator is bounded below by the lower-right sub-matrix of the inverse of the Bayesian information matrix, $J_{\text{CR}}^{-1}(k)$ [30], i.e.,

$$\text{Var}(\hat{\rho}_k) \geq J_{\text{CR}}^{-1}(k) = [J_{\text{CR}}^{-1}(k)]_{kk} \quad (75)$$

where the Bayesian information matrix is given by

$$\begin{aligned} [J_{\text{CR}}(k)]_{ij} &\triangleq \mathbb{E} \left[\frac{\partial \log f(\mathbf{Z}_k, \boldsymbol{\rho}_k)}{\partial \rho_i} \frac{\partial \log f(\mathbf{Z}_k, \boldsymbol{\rho}_k)}{\partial \rho_j} \right] \\ &= -\mathbb{E} \left[\frac{\partial^2 \log f(\mathbf{Z}_k, \boldsymbol{\rho}_k)}{\partial \rho_i \partial \rho_j} \right] \end{aligned}$$

where the expectation is taken with respect to the joint pdf and

$$\begin{aligned} \mathbf{Z}_k &\triangleq [Z_1, \dots, Z_k]^T \\ \boldsymbol{\rho}_k &\triangleq [\rho_0, \rho_1, \dots, \rho_k]^T \\ f(Z_k | \rho_k) &\propto \exp(\eta(Z_k)\rho_k - \phi_k(\rho_k)). \end{aligned} \quad (76)$$

It is assumed that the parameter ρ_k evolves through a Gauss-Markov model given by

$$f(\rho_k | \rho_{k-1}) = \frac{1}{\sqrt{2\pi}\sigma^2} \exp \left(-\frac{(\rho_k - \rho_{k-1})^2}{2\sigma^2} \right). \quad (77)$$

A recursive formula to evaluate the Bayesian sub-matrix, derived in [27], is given by

$$\begin{aligned} J_{\text{CR}}(k+1) &= -E_{\text{CR}}^{(2)}(k) \left(J_{\text{CR}}(k) + E_{\text{CR}}^{(1)}(k) \right)^{-1} E_{\text{CR}}^{(2)}(k) \\ &\quad + E_{\text{CR}}^{(3A)}(k) + E_{\text{CR}}^{(3B)}(k) \end{aligned} \quad (78)$$

where

$$\begin{aligned} E_{\text{CR}}^{(1)}(k) &\triangleq \mathbb{E} \left[-\frac{\partial^2}{\partial \rho_k^2} \log f(\rho_{k+1}|\rho_k) \right] \\ E_{\text{CR}}^{(2)}(k) &\triangleq \mathbb{E} \left[-\frac{\partial^2}{\partial \rho_k \partial \rho_{k+1}} \log f(\rho_{k+1}|\rho_k) \right] \\ E_{\text{CR}}^{(3A)}(k) &\triangleq \mathbb{E} \left[-\frac{\partial^2}{\partial \rho_{k+1}^2} \log f(\rho_{k+1}|\rho_k) \right] \\ E_{\text{CR}}^{(3B)}(k) &\triangleq \mathbb{E} \left[-\frac{\partial^2}{\partial \rho_{k+1}^2} \log f(Z_{k+1}|\rho_{k+1}) \right] \end{aligned}$$

and the expectation is again with respect to the joint pdf.

Theorem 6: For the Bayesian framework in (76) and (77), the recursive Bayesian information matrix in (78) is given by

$$J_{\text{CR}}(k+1) = (\sigma^2 + J_{\text{CR}}^{-1}(k))^{-1} + \sigma_{\eta_k}^2 \quad (79)$$

with $J_{\text{CR}}(0) = 0$.

Proof: For the density functions, $f(\rho_k|\rho_{k-1})$ and $f(Z_k|\rho_k)$ in (76), it can be verified that

$$E_{\text{CR}}^{(1)}(k) = \frac{1}{\sigma^2}, \quad E_{\text{CR}}^{(2)}(k) = -\frac{1}{\sigma^2}, \quad E_{\text{CR}}^{(3A)}(k) = \frac{1}{\sigma^2}$$

and

$$\begin{aligned} E_{\text{CR}}^{(3B)}(k) &= \int \int \frac{\partial^2 \phi_k(\rho_{k+1})}{\partial \rho_{k+1}^2} f(\rho_{k+1}, Z_{k+1}) d\rho_{k+1} dZ_{k+1} \\ &= \frac{\partial^2 \phi_k(\rho_{k+1})}{\partial \rho_{k+1}^2} = \sigma_{\eta_k}^2. \end{aligned}$$

The proof follows by plugging these quantities in (78). ■

D. Bayesian Chapman-Robbins Bound

A Bayesian version of the Chapman-Robbins bound (BCHRB) can be used to provide a lower bound on the variance of an estimator of ρ_k when there are no regularity assumptions on the likelihood. In fact, unlike the BCRB, the BCHRB can be evaluated for constrained likelihood functions where the domain of the likelihood is dependent on the unknown parameter.

BCHRB states that the variance of an estimator $\hat{\rho}_k$ of ρ_k is lower bounded as

$$\text{Var}(\hat{\rho}_k) - [T_k(\mathbf{h}_k) - 1]^{-1} \mathbf{h}_k \mathbf{h}_k^T \succeq \mathbf{0}$$

with \succeq in the positive semi-definite sense, where

$$T_k(\mathbf{h}_k) \triangleq \mathbb{E} \left[\left(\frac{f(\mathbf{Z}_k, \rho_k + \mathbf{h}_k)}{f(\mathbf{Z}_k, \rho_k)} \right)^2 \right],$$

and $\mathbf{h}_k \triangleq [0, h_1, \dots, h_k]^T$.

Theorem 7: The BCHRB for the parameter ρ_k can be expressed as

$$\text{Var}(\hat{\rho}_k) \geq \frac{1}{J_{\text{CH},k}}$$

where

$$J_{\text{CH},k} = \inf_{\mathbf{h}_k} \frac{T_k(\mathbf{h}_k) - 1}{h_k^2}$$

and

$$\begin{aligned} T_k(\mathbf{h}_k) &= \left(\prod_{j=1}^k M_{\eta}^{-2}(h_j) M_{\eta}(2h_j) \right) \times \\ &\exp \left[\frac{1}{\sigma^2} \sum_{j=1}^k (h_j - h_{j-1})^2 \right]. \end{aligned} \quad (80)$$

Proof: See Appendix B for details. ■

Remark 4: Since the performance bounds are derived for arbitrary exponential family distributions, they can also prove useful in a broad sense in classical as well as Bayesian parameter estimation.

E. Relation to Clock Offset Estimation

The performance bounds derived in the preceding subsections can be used to lower bound the mean square error (MSE) of the clock offset θ . Notice the similarity between the unconstrained and constrained likelihood functions for ξ and ψ in (5)-(8) and the general exponential family likelihood function considered in (68) and (69). Therefore, the bounds derived above for ρ are also applicable to the parameter ξ (and also ψ). The MSE of the clock offset θ can, in turn, be lower bounded using the bounds on ξ and ψ .

Using (4), the following result is immediate.

Proposition 1: The MSE of any estimator of θ can be expressed as

$$\text{MSE}(\hat{\theta}) = \frac{1}{4} \left(\text{Var}(\hat{\xi}) + \text{Var}(\hat{\psi}) \right) + \frac{1}{4} (b_{\xi} - b_{\psi})^2$$

where b_{ξ} and b_{ψ} are the biases of the estimators $\hat{\xi}$ and $\hat{\psi}$, respectively.

The explicit lower bounds on the MSE of any estimator $\hat{\theta}$ for classical as well as Bayesian framework in case of Gaussian and exponentially distributed likelihood functions can be evaluated as shown below.

1) *Gaussian Distribution - CRB:* If the likelihood function for ξ is Gaussian distributed (19), then using (20) and (71), it is straightforward to see that the CRB for any unbiased estimator $\hat{\xi}$ is given by

$$\text{Var}(\hat{\xi}) \geq \frac{\sigma_{\xi}^2}{N},$$

and a similar expression is applicable to $\hat{\psi}$ as well. Using Proposition 1, it can be concluded that

$$\text{MSE}(\hat{\theta}) \geq \frac{\sigma_{\xi}^2 + \sigma_{\psi}^2}{4N}. \quad (81)$$

As a remark, it is evident in this case that $\hat{\theta}_{\text{ML}}$ (22) is efficient in the sense that its MSE achieves (81) with equality (cf. Appendix C-A).

2) *Exponential Distribution - CHRB*: If the likelihood for ξ is exponentially distributed (31), using (11) and (32), it can be easily verified that

$$M_{\eta_\xi(U)}(h) = 1$$

and (74) becomes

$$\zeta(h) = \exp(\lambda_\xi h),$$

so that the statement of the CHRB (73) can be rewritten as

$$\text{Var}(\hat{\xi}) \geq \left[\inf_h \frac{\exp(\lambda_\xi h N) - 1}{h^2} \right]^{-1} = \frac{0.6476}{\lambda_\xi^2 N^2}$$

and similarly for $\hat{\psi}$. Using Proposition 1, it follows that

$$\begin{aligned} \text{MSE}(\hat{\theta}) &= \frac{1}{4} \left(\text{Var}(\hat{\xi}) + \text{Var}(\hat{\psi}) \right) + \frac{1}{4} (b_\xi - b_\psi)^2 \\ &\geq \frac{0.162}{N^2} \left(\frac{1}{\lambda_\xi^2} + \frac{1}{\lambda_\psi^2} \right) + \frac{1}{4} (b_\xi - b_\psi)^2. \end{aligned} \quad (82)$$

3) *Gaussian Distribution - BCRB*: In the Bayesian regime, if the likelihood function for ξ is Gaussian distributed (57), by using (58) and (79), it can be seen that

$$J_{\text{CR},\xi}(k+1) = \left(\sigma^2 + J_{\text{CR},\xi}^{-1}(k) \right)^{-1} + \frac{1}{\sigma_{\xi,k}^2},$$

with $J_{\text{CR},\xi}(0) = 0$. A similar line of reasoning can be followed to derive an analogous recursion for $J_{\text{CR},\psi}(k)$. The MSE of θ can be now be lower bounded as

$$\text{Var}(\hat{\theta}_k) \geq \frac{1}{4} \left(\frac{1}{J_{\text{CR},\xi}(k)} + \frac{1}{J_{\text{CR},\psi}(k)} \right). \quad (83)$$

4) *Exponential Distribution - BCHRB*: If the likelihood for ξ_k is exponentially distributed (63), (80) turns out to be

$$T_k(\mathbf{h}_k) = \exp \left(\lambda_\xi \sum_{j=1}^N h_j \right) \exp \left[\frac{1}{\sigma^2} \sum_{j=1}^k (h_j - h_{j-1})^2 \right].$$

In fact, we just have to notice that $\phi_\xi(\xi_k)$ is a constant function over ξ_k and $\eta_\xi(U_j) = \lambda_\xi$, so that (85) becomes

$$\mathbb{E}[\exp(2h_j \eta_\xi(U_j))] = \exp(\lambda_\xi h_j)$$

therefore $S(\mathbf{h}_k) = \exp \left(\lambda_\xi \sum_{j=1}^N h_j \right)$.

VI. SIMULATION RESULTS

This section aims to corroborate the theoretical results in preceding sections by conducting simulation studies in various scenarios. The performance of both the classical as well as Bayesian estimators is to be investigated. The measure of fidelity used to rate this performance is the MSE of the estimators for θ and θ_N . The parameter choice is $\sigma_\xi = \sigma_\psi = 0.1$ for both Gaussian and log-normally distributed likelihoods, while $\lambda_\xi = \lambda_\psi = 10$ for exponentially distributed likelihood functions.

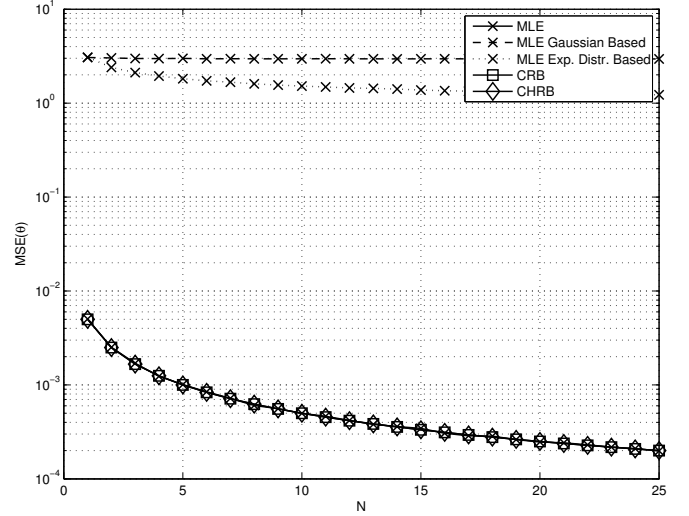


Fig. 3. MSE and bounds for estimating θ by using the MLE with log-normal likelihood.

A. Log-normal Distribution

No solution is reported thus far in literature in case the likelihood is log-normally distributed. The proposed estimators, $\hat{\theta}_{\text{ML}}$ (24) and $\hat{\theta}_N$ (62), can be used to determine an estimate for the clock offset in classical and Bayesian framework, respectively, as shown below.

1) *Classical Estimation Framework*: The existing approaches in literature only consider the Gaussian and the exponential cases, therefore (24) is a new result in the state-of-the-art about clock offset estimation. Fig. 3 shows a comparison between the proposed ML estimator (MLE) (24) in case of a log-normally distributed likelihood (23) with MLEs which (wrongly) assume that the likelihood is Gaussian and exponentially distributed, respectively. The plot shows that the latter approaches are not robust with respect to the likelihood distribution, and their performance is extremely poor if their assumptions do not hold. In addition, Fig. 3 also shows that the proposed MLE (24) is efficient since it attains the CRB (as well as the CHRB).

2) *Bayesian Estimation Framework*: Fig. 4 plots the MSE performance of the FGE (62) as well as the BCRB and BCHRB when the likelihoods are log-normally distributed (61), and $\sigma = 10^{-4}$. Firstly, it can be seen that the MSE of the proposed FGE coincides with the estimation bounds. Secondly, as in the classical estimation case, if we were to (wrongly) assume a Gaussian or exponential distribution for the likelihoods (38), the resulting FGEs would perform poorly, a fact is evident in Fig. 4 by observing the unboundedness and unpredictability of the dashed curve (Gaussian assumption for the likelihoods) and the dotted curve (likelihoods assumed exponentially distributed). This clearly establishes that the FGE (62), obtained assuming log-normally distributed likelihoods, allows a strong performance improvement with respect to existing estimators if the likelihood functions (38) are actually log-normally distributed.

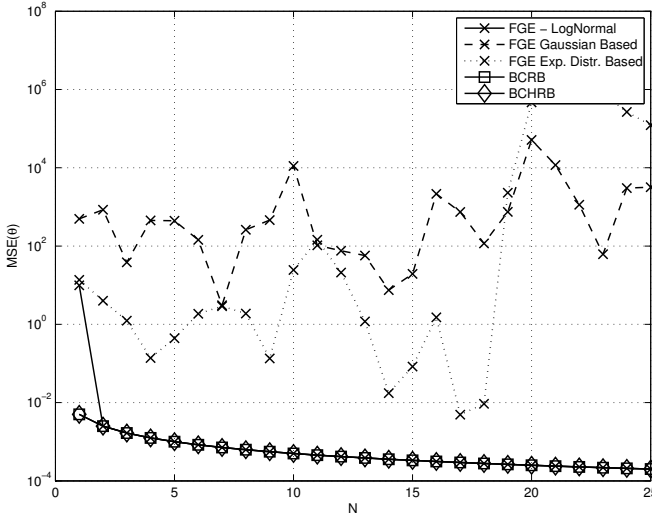


Fig. 4. MSE and bounds for estimating θ_N by using FGE with log-normal likelihood.

B. Estimator Performance vs Estimation Bounds

It will also be useful to assess the performance of the MLEs in Gaussian and exponential cases derived in Section III against the various benchmark estimation bounds derived in Section V. Similarly, the FGEs for Gaussian and exponential distributions, proposed in Section IV, can also be compared with the Bayesian bounds to study their MSE performance.

1) *Classical Estimation Framework*: Fig. 5 shows the performance comparison between the MSE of the MLEs (22) and (34) for Gaussian and exponentially distributed likelihood functions against the CRB and the CHRb, respectively. Firstly, it is evident that in the case of Gaussian distribution, the CRB and the CHRb coincide. Moreover, the MSE of $\hat{\theta}_{ML}$ also coincides with the aforementioned bounds. On the other hand, for an exponentially distributed likelihood function, due to its lack of regularity, the CRB cannot be derived, thus only the CHRb is shown. It can be observed that the MSE of $\hat{\theta}_{ML}$ is fairly close to CHRb, even though it does not coincide with it. From Fig. 5 the MSE of the MLEs for the Gaussian and exponential distribution case can be also compared. In order to ensure a fair comparison, parameters are chosen in a way to have the same variance of the observations for both distributions. From the MSE curves, one can infer that the MSE in case of an exponentially distributed likelihood is lower than the one for a Gaussian distribution as the number of observations N increases. This behavior is expected since it can be verified by the MSE expressions (86) and (87), reported in Appendix C, that in case of a Gaussian distribution, the MSE decreases proportionally to $1/N$, while in the exponential distribution case it decreases proportionally to $1/N^2$.

2) *Bayesian Estimation Framework*: In Fig. 6, the MSE performance of the FGEs $\hat{\theta}_N$ (60) and (67) is compared with BCRB and BCHRB for $\sigma = 10^{-4}$. As in the classical estimation scenario, it is evident that for Gaussian distributed likelihoods, the MSE using (60) for $\hat{\theta}_N$ coincides with the reported bounds. The MSE of the FGE derived assuming

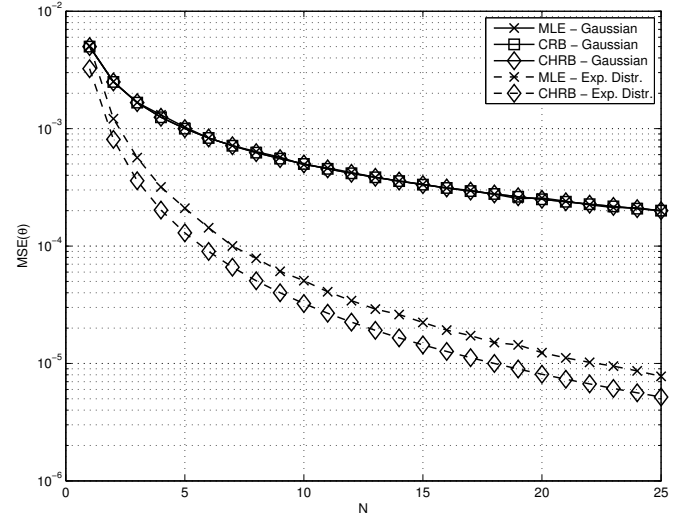


Fig. 5. MSE and bounds for estimating θ by using the MLE with Gaussian and exponentially distributed likelihood.

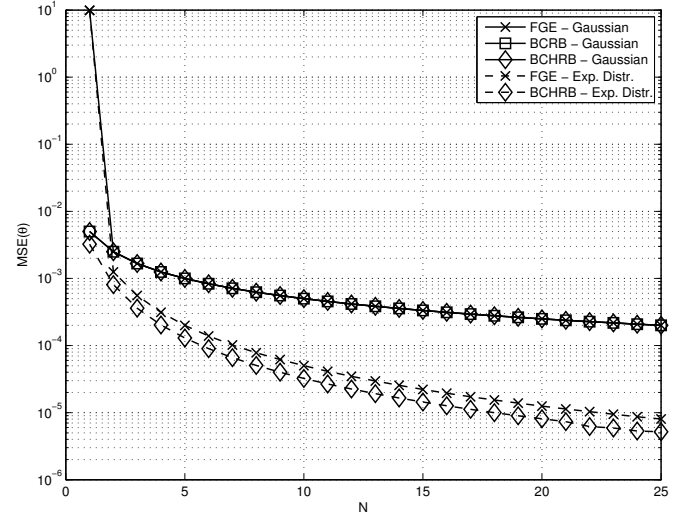


Fig. 6. MSE and bounds for estimating θ_N by using FGE with Gaussian and exponentially distributed likelihood.

exponentially distributed likelihoods (67) is plotted against the BCHRB as well in Fig. 6. It is clear that the MSE is quite close to BCHRB, although not coinciding with it, as exactly was the case in the classical estimation framework.

C. Comparing Classical and Bayesian frameworks

The estimators proposed in the classical and the Bayesian framework can also be compared with each other based on their MSE performance as the system noise decreases. The aim here is to show that the latter approaches the former as $\sigma \rightarrow 0$.

Fig. 7 depicts the MSE for the cases of Gaussian, exponential and log-normal distribution for the likelihoods with $N = 25$. In the plot, the horizontal lines represent the MSEs in the classical framework, obtained with the MLEs, as shown in (86) and (87) in Appendix C. It can be observed that, for all the three considered distributions, the MSE obtained by using

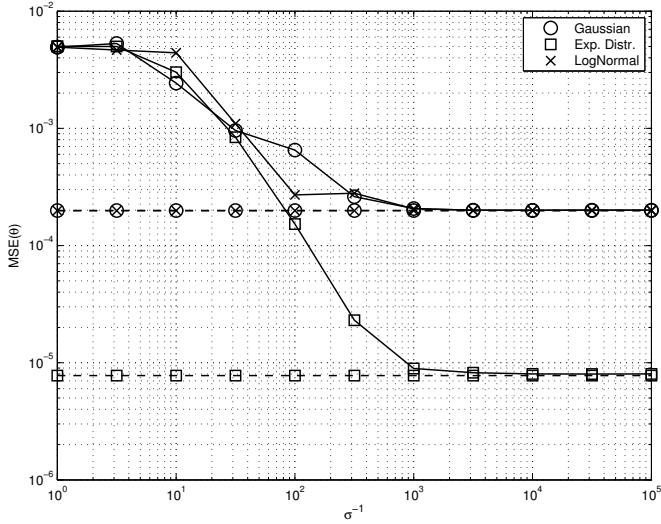


Fig. 7. MSE in estimating of θ_N vs σ . Solid lines: FGE; dashed lines: MLE (classical framework).

the FGE for estimating θ approaches the MSE of the MLEs as $\sigma < 10^{-3}$.

VII. CONCLUSIONS AND FUTURE WORK

The clock synchronization problem in sensor networks has received keen attention recently owing to its central role in critical network operations as duty cycling, data fusion, and node localization. Based on a two-way timing message exchange scenario, this work proposes a unified framework for the clock offset estimation problem when the likelihood function of the observation time stamps is Gaussian, exponential and log-normally distributed. A convex optimization based ML estimation approach is presented for clock offsets. The results known thus far for Gaussian and exponentially distributed network delays are subsumed in the general approach while the ML estimator is derived when the likelihood function is log-normally distributed. In order to study the case of a possibly time-varying clock offset, a Bayesian approach is also studied using factor graphs. The novel message passing strategy results in a closed form solution of the time-varying clock offset estimation problem. In order to compare various estimators, several lower bounds on the variance of an estimator have been derived in the classical as well as the Bayesian regime for likelihood functions which are arbitrary members of the exponential family, a wide class containing several distributions of interest. The theoretical findings are corroborated by simulation studies conducted in various scenarios.

In future, it will be useful to incorporate the effect of clock skew in the clock offset estimation model. This can result in further reduction of the re-synchronization periods. In addition, the results about pairwise synchronization can be used to build a framework for network-wide synchronization across a sensor network.

APPENDIX A PROOF OF THEOREM 5

The ratio of the likelihood functions can be expressed as

$$\begin{aligned} & \frac{f(\mathbf{Z}; \rho + h)}{f(\mathbf{Z}; \rho)} \\ &= \frac{e^{((\rho+h) \sum_{j=1}^N \eta(Z_j) - N\phi(\rho+h))} \prod_{j=1}^N \mathbb{I}(Z_j - \rho - h)}{e^{(\rho \sum_{j=1}^N \eta(Z_j) - N\phi(\rho))} \prod_{j=1}^N \mathbb{I}(Z_j - \rho)} \\ &= e^{(h \sum_{j=1}^N \eta(Z_j) - N\phi(\rho+h) + N\phi(\rho))} \prod_{j=1}^N \mathbb{I}(Z_j - \rho - h) \\ &= e^{(h \sum_{j=1}^N \eta(Z_j))} e^{(-N(\phi(\rho+h) + \phi(\rho)))} \prod_{j=1}^N \mathbb{I}(Z_j - \rho - h). \end{aligned}$$

The expectation of the ratio of the likelihood functions can now be calculated as

$$\begin{aligned} & \mathbb{E} \left(\frac{f(\mathbf{Z}; \rho + h)}{f(\mathbf{Z}; \rho)} \right)^2 \\ &= \mathbb{E} \left[e^{(2h \sum_{j=1}^N \eta(Z_j))} e^{(-2N(\phi(\rho+h) + \phi(\rho)))} \prod_{j=1}^N \mathbb{I}(Z_j - \rho - h) \right] \\ &= e^{(-2N(\phi(\rho+h) + \phi(\rho)))} \mathbb{E} \left[e^{(2h \sum_{j=1}^N \eta(Z_j))} \prod_{j=1}^N \mathbb{I}(Z_j - \rho - h) \right] \\ &= (M_{\eta(Z)}(h))^{-2N} \mathbb{E} \left[e^{(2h \sum_{j=1}^N \eta(Z_j))} \prod_{j=1}^N \mathbb{I}(Z_j - \rho - h) \right] \end{aligned}$$

where it follows from (11) that

$$(M_{\eta(Z)}(h))^{-2N} = e^{-2N(\phi(\rho+h) - \phi(\rho))}.$$

Since the samples Z_j are *i.i.d.*,

$$\begin{aligned} & \mathbb{E} \left[e^{(2h \sum_{j=1}^N \eta(Z_j))} \prod_{j=1}^N \mathbb{I}(Z_j - \rho - h) \right] \\ &= \left(\mathbb{E} \left[e^{(2h \eta(Z_j))} \mathbb{I}(Z_j - \rho - h) \right] \right)^N. \end{aligned}$$

With $\zeta(\cdot)$ defined in the theorem, the proof is complete.

APPENDIX B PROOF OF THEOREM 7

We have

$$\begin{aligned} T_k(\mathbf{h}_k) &\triangleq \mathbb{E} \left[\left(\frac{f(\mathbf{Z}_k, \boldsymbol{\rho}_k + \mathbf{h}_k)}{f(\mathbf{Z}_k, \boldsymbol{\rho}_k)} \right)^2 \right] \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(\frac{f(\mathbf{Z}_k, \boldsymbol{\rho}_k + \mathbf{h}_k)}{f(\mathbf{Z}_k, \boldsymbol{\rho}_k)} \right)^2 f(\mathbf{Z}_k, \boldsymbol{\rho}_k) d\mathbf{Z}_k d\boldsymbol{\rho}_k \\ &= S(\mathbf{h}_k) \int_{-\infty}^{+\infty} \frac{f(\boldsymbol{\rho}_k + \mathbf{h}_k)^2}{f(\boldsymbol{\rho}_k)} d\boldsymbol{\rho}_k \end{aligned}$$

where

$$S(\mathbf{h}_k) \triangleq \int_{-\infty}^{+\infty} \left(\frac{f(\mathbf{Z}_k | \boldsymbol{\rho}_k + \mathbf{h}_k)}{f(\mathbf{Z}_k | \boldsymbol{\rho}_k)} \right)^2 f(\mathbf{Z}_k | \boldsymbol{\rho}_k) d\mathbf{Z}_k. \quad (84)$$

Continuing with the calculations

$$\begin{aligned}
T_h(\mathbf{h}_k) &= S(\mathbf{h}_k) \int_{-\infty}^{+\infty} \frac{f(\boldsymbol{\rho}_k + \mathbf{h}_k)^2}{f(\boldsymbol{\rho}_k)} d\boldsymbol{\rho}_k \\
&= S(\mathbf{h}_k) \int_{-\infty}^{+\infty} \frac{f^2(\rho_0 + h_0)}{f(\rho_0)} \times \\
&\quad \prod_{j=1}^k \frac{f^2(\rho_j + h_j | \rho_{j-1} + h_{j-1})}{f(\rho_j | \rho_{j-1})} d\boldsymbol{\rho}_k \\
&= S(\mathbf{h}_k) \int_{-\infty}^{+\infty} \prod_{j=1}^k \frac{f^2(\rho_j + h_j | \rho_{j-1} + h_{j-1})}{f(\rho_j | \rho_{j-1})} d\boldsymbol{\rho}_k.
\end{aligned}$$

Since $\frac{f^2(\rho_j + h_j | \rho_{j-1} + h_{j-1})}{f(\rho_j | \rho_{j-1})}$ can be verified to be equal to $(j = 1, \dots, k)$

$$\begin{aligned}
&\frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{\rho_{j-1}^2}{2\sigma^2} + \frac{\rho_j + 2(h_j - h_{j-1})}{\sigma^2} \rho_{j-1} \right] \times \\
&\exp \left[-\frac{(h_j - h_{j-1})^2}{\sigma^2} \right] \exp \left[-\frac{\rho_j^2}{2\sigma^2} - \frac{2\rho_j(h_j - h_{j-1})}{\sigma^2} \right]
\end{aligned}$$

it turns out that

$$\int_{-\infty}^{+\infty} \frac{f^2(\rho_j + h_j | \rho_{j-1} + h_{j-1})}{f(\rho_j | \rho_{j-1})} d\rho_{j-1} = \exp \left[\frac{(h_j - h_{j-1})^2}{\sigma^2} \right]$$

that brings to

$$T_k(\mathbf{h}_k) = S(\mathbf{h}_k) \exp \left[\frac{1}{\sigma^2} \sum_{j=1}^k (h_j - h_{j-1})^2 \right].$$

It can be easily verified that (84) can be written as

$$S(\mathbf{h}_k) = \prod_{j=1}^k \int_{-\infty}^{+\infty} \left(\frac{f(Z_j | \rho_j + h_j)}{f(Z_j | \rho_j)} \right)^2 f(Z_j | \rho_j) dZ_j.$$

Moreover, it can be noted that

$$\begin{aligned}
\left(\frac{f(Z_j | \rho_j + h_j)}{f(Z_j | \rho_j)} \right)^2 &= \exp [-2(\phi_\rho(\rho_j + h_j) - \phi_\rho(\rho_j))] \times \\
&\quad \exp (2h_j \eta_\rho(Z_j))
\end{aligned}$$

and therefore

$$\begin{aligned}
\int_{-\infty}^{+\infty} \left(\frac{f(Z_j | \rho_j + h_j)}{f(Z_j | \rho_j)} \right)^2 f(Z_j | \rho_j) dZ_j &= M_{\eta_\rho}^{-2}(h_j) \times \\
&\quad \mathbb{E} [\exp (2h_j \eta_\rho(Z_j))] .
\end{aligned}$$

Then, since

$$\mathbb{E} [\exp (2h_j \eta_\rho(Z_j))] = \exp (\phi_\rho(\rho_j + 2h_j) - \phi_\rho(\rho_j)) \quad (85)$$

it can be easily seen that

$$\int_{-\infty}^{+\infty} \left(\frac{f(Z_j | \rho_j + h_j)}{f(Z_j | \rho_j)} \right)^2 f(Z_j | \rho_j) dZ_j = M_{\eta_\rho}^{-2}(h_j) M_{\eta_\rho}(2h_j),$$

thus getting

$$S(\mathbf{h}_k) = \prod_{j=1}^k M_{\eta_\rho}^{-2}(h_j) M_{\eta_\rho}(2h_j).$$

APPENDIX C

MSE EXPRESSIONS FOR ML ESTIMATORS

A. Gaussian Distribution

If the likelihood for ξ is Gaussian distributed (19), the MLE is given by (21). Since the variance of the readings U_j is σ_ξ^2 and the MLE (21) is unbiased, it is straightforward to see that

$$\text{MSE}(\hat{\xi}_{\text{ML}}) = \frac{\sigma_\xi^2}{N},$$

and similarly for $\hat{\psi}_{\text{ML}}$. Given (18) and Proposition 1, it can be concluded that

$$\text{MSE}(\hat{\theta}_{\text{ML}}) = \frac{\sigma_\xi^2 + \sigma_\psi^2}{4N}. \quad (86)$$

B. Exponential Distribution

If the likelihood for ξ is exponential distributed (31), the MLE is given by (33). Through simple algebra it can be seen that $U_{(1)}$ is exponentially distributed with parameter $\lambda'_\xi = \lambda_\xi N$, so that $\text{Var}(\hat{\xi}_{\text{ML}}) = \frac{1}{\lambda_\xi^2 N^2}$. It can be noticed that $\hat{\xi}_{\text{ML}}$ is a biased estimator for ξ , with bias $b_{\xi, \text{ML}} \triangleq \frac{1}{\lambda_\xi N}$. Similarly, $\text{Var}(\hat{\psi}_{\text{ML}}) = \frac{1}{\lambda_\psi^2 N^2}$ and $b_{\psi, \text{ML}} \triangleq \frac{1}{\lambda_\psi N}$. Therefore, given (18) and Proposition 1, it can be concluded that

$$\text{MSE}(\hat{\theta}_{\text{ML}}) = \frac{0.25}{N^2} \left(\frac{1}{\lambda_\xi^2} + \frac{1}{\lambda_\psi^2} \right) + \frac{0.25}{N^2} \left(\frac{1}{\lambda_\xi} - \frac{1}{\lambda_\psi} \right)^2. \quad (87)$$

REFERENCES

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102-114, Aug. 2002.
- [2] B. Sundararaman, U. Buy, and A. D. Kshemkalyani, "Clock synchronization for wireless sensor networks: A survey," *Ad-Hoc Netw.*, vol. 3, no. 3, pp. 281-323, Mar. 2005.
- [3] B. Sadler and A. Swami, "Synchronization in sensor networks: An overview," in *Proc. IEEE Military Commun. Conf. (MILCOM 2006)*, pp. 1-6, Oct. 2006.
- [4] M. Maroti, B. Kusy, G. Simon, and A. Ledeczi, "The flooding time synchronization protocol," in *Proc. 2nd Int. Conf. Embedded Networked Sensor Systems*, ACM Press, pp. 39-49, Nov. 2.
- [5] S. Ganeriwal, R. Kumar, and M. B. Srivastava, "Timing-sync protocol for sensor networks," in *Proc. SenSys*, Los Angeles, CA, pp. 138-149, Nov. 2003.004.
- [6] J. Elson, L. Girod, and D. Estrin, "Fine-grained network time synchronization using reference broadcasts," in *Proc. 5th Symp. Operating System Design Implementation*, Boston, MA, pp. 147-163, Dec. 2002.
- [7] C. Bovy, H. Mertodimedjo, G. Hooghiemstra, U. Uijterwaal, and P. Miegheem, "Analysis of end-to-end delay measurements in internet," in *Proc. Passive Active Meas. Workshop*, Fort Collins, CO, pp. 26-33, Mar. 2002.
- [8] M. D. Lemmon, J. Ganguly, and L. Xia, "Model-based clock synchronization in networks with drifting clocks," in *Proc. Int. Symp. Dependable Computing*, Washington, DC, pp. 177-184, 2000.
- [9] O. Simeone and U. Spagnolini, "Distributed time synchronization in wireless sensor networks with coupled discrete-time oscillators," *EURASIP Jour. Wireless Commun. and Networking*, 2007.
- [10] D. Zennaro, E. Dall'Anese, T. Erseghe, and L. Vangelista, "Fast clock synchronization in wireless sensor networks via ADMM-based consensus," in *Proc. of 9th Int. Symp. Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, Princeton, NJ, May 9-13, 2011.
- [11] Y.-W. Hong and A. Scaglione, "A scalable synchronization protocol for large scale sensor networks and its applications," *IEEE Jour. Selected Areas Commun.*, vol. 23, no. 5, pp. 1085-1099, May 2005.

- [12] R. Pagliari and A. Scaglione, "Scalable network synchronization with pulse-coupled oscillators," *IEEE Trans. Mob. Computing*, vol. 10, no. 3, pp. 392-405, Mar. 2011.
- [13] H. S. Abdel-Ghaffar, "Analysis of synchronization algorithm with time-out control over networks with exponentially symmetric delays," *IEEE Trans. Commun.*, vol. 50, no. 10, pp. 1652-1661, Oct. 2002.
- [14] D. R. Jeske, "On the maximum likelihood estimation of clock offset," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 53-54, Jan. 2005.
- [15] K.-L. Noh, Q. M. Chaudhari, E. Serpedin, and B. Suter, "Novel clock phase offset and skew estimation using two-way timing message exchanges for wireless sensor networks," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 766-777, Apr. 2007.
- [16] Q. M. Chaudhari, E. Serpedin, and K. Qaraqe, "On maximum likelihood estimation of clock offset and skew in networks with exponential delay," *IEEE Trans. Sig. Processing*, vol. 56, no. 4, pp. 1685-1697, Apr. 2008.
- [17] M. Leng and Y.-C. Wu, "On clock synchronization algorithms for wireless sensor networks under unknown delay," *IEEE Trans. Veh. Tech.*, vol. 59, no. 1, pp. 182-190, Jan. 2010.
- [18] G. Xiong and S. Kishore, "Analysis of distributed consensus time synchronization with Gaussian delay over wireless sensor networks," *EURASIP Jour. Wireless Commun. and Networking*, 2009.
- [19] G. Xiong and S. Kishore, "Discrete-time second-order distributed consensus time synchronization algorithm for wireless sensor networks," *EURASIP Jour. Wireless Commun. and Networking*, 2009.
- [20] Q. M. Chaudhari and E. Serpedin, "On minimum variance unbiased estimation of clock offset in a sender-receiver protocol," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2893-2904, June 2010.
- [21] A. Ahmad, A. Noor, E. Serpedin, H. Nounou, and M. Nounou, "On clock offset estimation in wireless sensor networks with Weibull distributed network delays," *Int. Conf. Patt. Recognition*, Istanbul, Turkey, Aug. 2010.
- [22] Y.-C. Wu, Q. M. Chaudhari, and E. Serpedin, "Clock synchronization of wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 28, no. 1, pp. 124-138, Jan. 2011.
- [23] N. M. Freris, S. R. Graham, and P. R. Kumar, "Fundamental limits on synchronizing clocks over networks," *IEEE Trans. Auto. Control* vol. 56, no. 6, pp. 1352-1364, June 2011.
- [24] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498-519, Feb. 2001.
- [25] M. J. Wainwright, T. S. Jaakola, and A. S. Willsky, "A new class of upper bounds on the log partition function," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2313-2335, July 2005.
- [26] D. G. Chapman and H. Robbins, "Minimum variance estimation without regularity assumptions," *Annals of Math. Statistics*, vol. 22, no. 4, pp. 581-586, Dec. 1951.
- [27] P. Tichavsky, C. H. Muravchik, and A. Nehorai, "Posterior Cramer-Rao bounds for discrete-time nonlinear filtering," *IEEE Trans. Signal Processing*, vol. 46, no. 5, pp. 1386-1396, May 1998.
- [28] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. I, Estimation Theory*, Englewood Cliffs, NJ: Prentice Hall, 1993.
- [29] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge, UK, Cambridge University Press, 2003.
- [30] H. L. Van Trees and K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, Wiley, 2007.